

EUROSLA MONOGRAPHS SERIES 3

GRAMMATICAL DEVELOPMENT IN SECOND LANGUAGES: Exploring the boundaries of Processability Theory

EDITED BY

CAMILLA BETTONI

University of Verona

BRUNO DI BIASE

University of Western Sydney

EUROPEAN SECOND LANGUAGE ASSOCIATION 2015

Eurosla monographs

Eurosla publishes a monographs series, available open access on the association's website. The series includes both single-author and edited volumes on any aspect of second language acquisition. Manuscripts are evaluated with a double blind peer review system to ensure that they meet the highest qualitative standards.

Series editor

Gabriele Pallotti,
University of Modena and Reggio Emilia

Assistant editors

Amanda Edmonds,
Université de Pau et des Pays de l'Adour

Kevin McManus,
University of York

Fabiana Rosi,
University of Modena and Reggio Emilia

© The Authors 2015

Published under the Creative Commons

“Attribution Non-Commercial No Derivatives 3.0” license

ISBN 978-1-329-42765-5

First published by Eurosla, 2015

Graphic design and layout: Pia 't Lam

An online version of this volume can be downloaded from eurosla.org

*We dedicate this volume to
Maria Agata Di Biase,
who would speak to anyone with
little recourse to a second language,
and to Noel Macainsh,
who would not speak in
his second languages unless
he could use a well phrased sentence.*

Editorial board

Cecilia Andorno, University of Pavia
Dalila Ayoun, University of Arizona
Camilla Bardel, Stockholm University
Alessandro Benati, University of Greenwich
Sandra Benazzo, Université Lille 3
Giuliano Bernini, University of Bergamo
Camilla Bettoni, University of Verona
Marina Chini, University of Pavia
Jean-Marc Dewaele, Birkbeck College, UCL
Anna Giacalone Ramat, University of Pavia
Roger Gilabert, University of Barcelona
Pedro Guijarro-Fuentes, Plymouth University
Gisela Håkansson, Lund University
Henriëtte Hendriks, University of Cambridge
Martin Howard, University College Cork
Gabriele Kasper, University of Hawai'i at Ma-noa
Judith Kormos, Lancaster University
Folkert Kuiken, University of Amsterdam
Maise Martin, University of Jyväskylä
James Milton, Swansea University
John Norris, Georgetown University
Lourdes Ortega, Georgetown University
Simona Pekarek-Doehler, Université de Neuchâtel
Manfred Pienemann, University of Paderborn
Leah Roberts, University of York
Jason Rothman, University of Iowa
Michael Sharwood Smith, Heriot-Watt University Edinburgh
Nina Spada, University of Toronto
Richard Towell, University of Salford
Danijela Trenkic, University of York
Ada Valentini, University of Bergamo
Ineke Vedder, University of Amsterdam
Christiane von Steutterheim, Heidelberg University
Johannes Wagner, University of Southern Denmark

Table of contents

Foreword <i>Camilla Bettoni and Bruno Di Biase</i>	7
Abbreviations	13
Part I	
Introducing and developing Processability Theory <i>Camilla Bettoni and Bruno Di Biase</i>	15
1 Processability Theory: theoretical bases and universal schedules <i>Camilla Bettoni and Bruno Di Biase</i>	19
Part II	
The developmental path across languages <i>Camilla Bettoni and Bruno Di Biase</i>	81
2 The development of English as a second language <i>Bruno Di Biase, Satomi Kawaguchi and Yumiko Yamaguchi</i>	85
3 The development of Italian as a second language <i>Bruno Di Biase and Camilla Bettoni</i>	117
4 The development of Japanese as a second language <i>Satomi Kawaguchi</i>	149
Part III	
Exploring new issues <i>Camilla Bettoni and Bruno Di Biase</i>	173
5 Acquiring case marking in Russian as a second language: an exploratory study on subject and object <i>Daniele Artoni and Marco Magnani</i>	177
6 The development of case in a bilingual context: Serbian in Australia <i>Bruno Di Biase, Camilla Bettoni and Lucija Medojević</i>	195

7	Exploring the acquisition of differential object marking (DOM) in Spanish as a second language	213
	<i>Bruno Di Biase and Barbara Hinger</i>	
8	The development of constituent questions in Italian as a second language	243
	<i>Camilla Bettoni and Giorgia Ginelli</i>	
9	Acquiring V2 in declarative sentences and constituent questions in German as a second language	259
	<i>Louise Jansen and Bruno Di Biase</i>	
10	Exploring Processability Theory-based hypotheses in the second language acquisition of a child with autism spectrum disorder	275
	<i>Tonya G. Agostini and Catherine T. Best</i>	
11	Connecting CALL and second language development: online tandem learning of Japanese	291
	<i>Satomi Kawaguchi</i>	
	References	307
	About the authors	333

Foreword

Processability Theory (PT from now on) is a psycholinguistic theory of second language acquisition (SLA from now on), first formulated in a book-length publication by Manfred Pienemann in 1998. Since then, PT has earned a place in many SLA introductions, handbooks, companions and encyclopedias such as Doughty & Long (2003), Gass & Mackey (2012), Kroll & de Groot (2005), Macaro (2013), Ortega (2009) and Robinson (2012). It is one of nine theories of SLA singled out and discussed by VanPatten & Williams (2007), along with four others sharing a cognitive processing approach. Like other theories, PT has its strengths and weaknesses, and does not deal with all the phenomena and processes constraining SLA or contributing to it. However, few theories seem to accommodate such a variety of phenomena or offer the basis for so many new developments. What PT offers is a principled transitional paradigm that deals specifically with grammatical development, and accounts for it. It also contributes an explicit and universal definition of developmental stages, which may be applied in principle to any language-specific developmental trajectory, and which in turn offers a stable point of reference for investigating typologically diverse L1-L2 constellations, learning modalities, environments and populations. From a practical point of view, PT can help in assessing language development in individual learners as well as in constructing a syllabus appropriate for their stage of development. In terms of new directions, as Jordan (2004: 227) remarked, PT “can be seen as ‘progressive’ [...] extending its domain, refining its concepts, making the variables more operational, attracting more research.”

More than fifteen years have passed since the publication of Pienemann’s book on PT in 1998; and before that, it took almost twenty years to mould into PT the initial achievements of the ZISA (*Zweitspracherwerb ausländischer Arbeiter*) Multidimensional Model (cf. foremost Meisel, Clahsen & Pienemann 1981). Over this period, not only has the whole field of SLA grown exponentially, but PT’s feeder disciplines have also advanced significantly. PT has paralleled this growth, and widened its scope in several directions. We very briefly retrace its history in order to place our volume within this developing context.

In the early eighties the ZISA team proposed a Multidimensional Model to account for their considerable body of data and, in particular, to explain the staged development of German word order. The two dimensions of the Model were, on the one hand, a psycholinguistic and potentially universal dimension, and on the

other, a socio-psychological one, no longer pursued. Within a broadly Chomskyan generative framework, the ZISA team guided by Jürgen Meisel provided the SLA field with two further fruitful contributions: first, while lively debates moved SLA more and more towards processing approaches (e.g., McLaughlin 1978, 1980 vs Krashen 1976, 1979), Harald Clahsen proposed an explanation for the apparently invariant stages observed in the acquisition of German word order in terms of cognitive strategies, the accumulation of rules determining the learner's grammatical progress (Clahsen 1980); and secondly, Manfred Pienemann successfully addressed the theory-practice link by focusing on the connection between development and teaching with his Teachability Hypothesis (Pienemann 1984, 1985, 1989), showing experimentally that it is not possible to alter the trajectory of acquisition through L2 teaching. This latter hypothesis, which has since remained unchanged, is still, in our opinion, the key to the popularity of Pienemann's approach, its potential already recognised in Larsen Freeman & Long (1991: 270-287).

When Manfred Pienemann moved to Australia in 1982, Malcolm Johnston's large structural English L2 data set (Johnston 1985a) provided the two researchers with the challenge to test out the assumed universality of ZISA's cognitive account. Assuming the universality of human cognition, if the constraints operative for the development of German L2 syntax are cognitively based, then they should turn out to be operative not only for German word order but also for other L2s and possibly other domains, such as morphology, provided the target structures meet the requirements of those processable by each particular strategy (Pienemann & Johnston 1986; 1987). In other words, this framework had a predictive potential, as Larsen-Freeman & Long (1991: 287) pointed out. Thus, taking on board Pienemann's (1984) teachability experiments and other contemporary works pointing in the same direction, the framework could predict for example that, no matter what the teacher may try, the learners would not 'skip' a stage of acquisition and could only learn what they are developmentally 'ready' to learn. The reader will quickly notice that, having specified what belongs to what stage, this prediction is falsifiable. Pienemann's Teachability Hypothesis however did not dismiss the role of instruction altogether. Rather it constrained its potential to improvements in the rate, not the route, of acquisition. This 'predictive framework' provided key theoretical insights, and generated publications by both Johnston and Pienemann on a range of issues of interest to SLA at the time (cf. Di Biase 2000), such as whether there is a 'natural' progression in learners of English from different L1 background (Johnston 1985b), what factors influence language development (Pienemann & Johnston 1987), how learners develop their own grammar (Johnston 1987), and what exactly is the influence of instruction on L2 processing (Pienemann 1987). An early result of the Pienemann and Johnston predictive framework's potential was the application, in association with Geoff Brindley, of the hypothesised developmental stages to the construction of a principled and interlanguage-sensitive lan-

guage assessment and testing procedure for ESL (Pienemann, Johnston & Brindley 1988), providing the foundation for Pienemann's later Rapid Profiling. The application of this innovative predictive framework to English, however, was not without problems, as pointed out by Larsen-Freeman & Long (1991: 275-ff), yet the schedules for the development of English morphosyntax elaborated in the eighties on Johnstons's (1985a, 1985b) cross-sectional data from 12 Polish and 12 Vietnamese immigrants to Australia have provided key empirical evidence ever since (cf., e.g., Pienemann 1998: 165-181; Pienemann 2011a: 8-11).

Over the nineties, the empirical base of PT widened in several directions. First, the structural results on German L2 and English L2 data were supported by positively testing the Steadiness Hypothesis, according to which task variation does not produce variation in the learner's procedural schedules (Pienemann, Mackey & Thornton 1991; Pienemann & Mackey 1993). Secondly, a fruitful connection was established between input and interaction on the one hand and development on the other in the area of question formation (Mackey 1999). Thirdly, the typological validation widened with initial work on the acquisition of Japanese (Doi & Yoshioka 1987, 1990; Huter 1996, 1997), and Arabic (Mansouri 1995, 1997, 1999). Thus, the limitations of the original strategy-based explanation became more evident, as much as the need for a more explicit approach to grammatical representation if the theory was to work cross-linguistically. PT came about when Pienemann (1998), abandoning the strategies approach, grounded his new theory's psychological plausibility on Levelt's 'Blueprint for the Speaker' (1989), an explicit model of language generation, itself a successful integration of several strands of psycholinguistic research focusing on speech processing. At the same time, following Levelt's (1989) own example, as well as Pinker's (1984), Pienemann (1998) introduced a second crucial innovation that provided PT with the necessary grammar-theoretical basis to test and support its typological plausibility, namely, Kaplan & Bresnan's (1982) Lexical Functional Grammar (LFG from now on). In Kaplan's (1995: 7) own words the "LFG formalism, which has evolved from previous computational, linguistic and psycholinguistic research, provides a simple set of devices for describing the common properties of all human languages and the particular properties of individual languages". PT then further broadened and consolidated its typological spread, moving from its German and English focus towards a greater variety of languages, such as Arabic (Mansouri 2002, 2005), Chinese (e.g., Zhang 2002, 2004), French (Ågren 2009), Italian and Japanese (e.g., Di Biase & Kawaguchi 2002; Kawaguchi 2005), as well as Swedish and other Scandinavian languages (e.g., Pienemann & Håkansson 1999; Glahn et al. 2001).

Further PT developments over the first decade of this century include the Developmentally Moderated Transfer Hypothesis (Pienemann, Di Biase, Kawaguchi & Håkansson 2005) to account for L1-L2 transfer, following on from work on typological proximity (Håkansson, Pienemann & Sayehli 2002). This

hypothesis claims that transfer from L1 to L2 can only happen if it is in accordance with PT's schedules, that is, in a nutshell, you can transfer only what you can process. Furthermore, the plausibility of the theory was extended to acquisition contexts other than classic adult L2 ones, such as bilingual language acquisition (e.g., Itani-Adams 2008), children with Specific Language Impairment (e.g., Håkansson 2001, 2005; Håkansson, Salameh & Nettelbladt 2003), child L2 learning (Yamaguchi 2010), and in the emergence of creole languages (Plag 2008a, 2008b, 2011). Finally, the range of applications of the original PT to language teaching and language testing has also expanded over the years, involving several new languages, teaching situations and ways of testing, while generating new hypotheses, such as the syllabus construction hypothesis (Pienemann & Kessler 2007), and the developmentally moderated feedback hypothesis (Di Biase 2002, 2008; Nuzzo & Bettoni 2011); as well as further work on question formation in English as foreign language contexts (Sakai 2008).

Some ten years ago, PT's framework (Pienemann 1998) widened substantially when Pienemann, Di Biase & Kawaguchi (2005) followed up on an idea first introduced in 2002 by Di Biase & Kawaguchi in coming to terms with their two nonconfigurational languages, Italian and Japanese respectively. By incorporating LFG's syntactically motivated discourse functions and Lexical Mapping Theory (Bresnan 2001; Dalrymple 2001; Falk 2001), PT added a new discourse-pragmatically motivated syntactic component to its 'classic' syntactically motivated morphological module. Since the formulation of this 2005 extension, its two main hypotheses, the Topic Hypothesis and the Lexical Mapping Hypothesis, have been tested on various languages (cf., e.g., Itani-Adams 2009; Zhang 2007; Yamaguchi 2010; and Bettoni & Di Biase 2011 for the former hypothesis; mainly Kawaguchi 2007, 2009a but also by Wang 2009, and Keatinge & Kessler 2009 for the latter hypothesis). However, these works have just begun to explore the two hypotheses, and more needs to be done.

Given the above historical sketch of PT, the aim of our volume is twofold. First, we intend to provide a consistent, if concise, new presentation of PT's main tenets, as we see them in connection with our proposal for a broader 'prominence hypothesis' to approach the syntax-discourse interface (cf. part I). We feel that a new presentation is necessary in order to clarify some theoretical and terminological issues. For instance, in Pienemann & Kessler (2011), some of the original transformational terminology from much earlier work (e.g., Pienemann, Johnston & Brindley 1988) is still used for describing English PT's schedules. While this may have some practical advantages in terms of continuity, we offer some proposals for updating the terminology so as to incorporate current LFG into PT's framework in a consistent fashion. For example, connecting LFG's discourse functions and its Lexical Mapping Hypothesis to L2 development enables us to investigate their operation analytically, and interpret the learner's ability, among others, to topicalise a

grammatical function other than the subject or choose between active and passive constructions as something other than a purely structural operation. In other words, the processing of discourse-pragmatic information does play a principled role, and after all LFG is a declaredly nonderivational grammatical theory. As Asudeh & Toivonen (2010: 454-455) put it, a basic principle of LFG theory is that “grammatical information grows *monotonically* (Bresnan 2001: ch. 5), i.e., in an information-preserving manner. [...] One general consequence is that there can be no destructive operations in syntax. For example, relation-changing operations, such as passive, cannot be syntactic, because that would require destructive remapping of grammatical functions” (emphasis in the original).

The second aim of our volume is to make a contribution to theory construction. For example, we take a more analytical approach to PT and look first at morphology separately from syntax, at declaratives separately from questions, and then at their interfaces. Furthermore, we propose to recast the 2005 Topic Hypothesis as the Prominence Hypothesis, and reformulate the Lexical Mapping Hypothesis. The changes in the staging of L2 development deriving from these proposals, broadly presented in part I, are then illustrated in part II through the recast learners’ schedules of three typologically distant and well tested languages in PT: English, Italian and Japanese. Furthermore, we intend to contribute to theory construction in PT by exploring new possibilities and providing a coherent context for current work – that is, new work which draws on the consequences of developments in PT’s feeder disciplines, and explores issues, languages and applications not previously treated in PT. Thus, in part III of our volume the scope of PT is widened in several directions. Among them, Russian is a new language for PT, which exemplifies the way in which grammatical case can, and must, integrate morphological and syntactic considerations. Another new area is the treatment of Differential Object Marking, as it emerges in Spanish L2, itself a language scantily examined within PT (but cf. Bonilla 2012, 2014). Constituent questions, an older SLA area, receive new treatment within our Prominence Hypothesis, thanks to Mycock’s (2007) pioneering work within the nonderivational LFG framework. For instance, our more analytical approach to PT enables a comparative treatment of questions and declaratives, leading to results for German L2 development that are theoretically interesting in the history of PT. Part III also includes explorations into whether PT’s stages hold also with autistic L2 learners, and concludes with a sally into technological innovations applied to language teaching and learning, and the role PT may play in this important area. Our hope is that the exploratory nature of some of this work will suggest potential lines of further development for both the theory itself and its applications – thus contributing to PT’s promise as a ‘progressive’ theory. Where the evidence for our new hypotheses is sometimes based on few learners – even one learner in longitudinal studies – we can only invite researchers to test our claims on richer data sets.

The principal audience for this volume consists of SLA researchers and graduate students, advanced undergraduate students and their instructors. The editors and authors assume little previous knowledge of PT on the part of the reader. On the other hand, we must warn readers that this is not a comprehensive introduction to PT, because our volume does not discuss important areas such as the Developmentally Moderated Transfer Hypothesis, learner variability or applications such as the Rapid Profiling. In addition to limitations of space in a single volume, these gaps may also be attributed to our wish to prioritise for PT fresh interpretations of earlier achievements, new ground, and new pointers for future research.

The ideas presented in this volume build on previous research on PT by numerous colleagues, relying foremost, of course, on Manfred Pienemann's (1998) cornerstone publication. This does not mean that the originator of PT agrees with all aspects of our interpretation of his theory or with all our contributing authors' new developments. His reservations have helped us clarify our own thoughts, and we thank him for sharing them with us. Care is taken throughout the volume by editors and authors alike to signal to the reader differences or potential disagreement between previous versions of PT and the explorations presented here, and to provide references to specific works whenever appropriate.

Among our many colleagues working in the PT framework, we wish to single out Gisela Håkansson, Yuki Itani-Adams, Junko Iwasaki, Louise Jensen, Satomi Kawaguchi, Jörg Keßler, Fethi Mansouri, Gabriele Pallotti, Manfred Pienemann, and Yanyin Zhang among the older generation; and Daniele Artoni, Marco Magnani, Lucija Medojevic, Elena Nuzzo, Jacopo Torregrossa, Yumiko Yamaguchi, Kenny Wang, and Karoline Wirbatz among the younger generation. We thank them all not only for their scholarly contribution but also for the personal friendships that have developed with them over the many PT meetings held in Australia and Europe in the last two decades or so. As editors, we wish to thank foremost Gabriele Pallotti and Manfred Pienemann, then, alphabetically, Cinzia Avesani, Cathi Best, Gisela Håkansson, Barbara Hinger, Louise Jansen, Jörg Keßler, Marco Magnani, Louise Mycock, Elena Nuzzo, Lourdes Ortega, Valeria Peretokina, Ingo Plag, Ruying Qi, Peter Robinson, Jason Shaw for their help and encouragement, and the anonymous reviewers for their thoughtful and constructive comments on the volume as a whole and/or single chapters. Needless to say, remaining errors are our responsibility.

30 May 2015

Camilla Bettoni
Bruno Di Biase

Abbreviations

1	1 st person
2	2 nd person
3	3 rd person
ABL	ablative
ABS	absolutive
ACC	accusative
ADJ	adjunct
ASP	aspect
a-structure	argument structure
AUX	auxiliary
BEN	benefactive
CALL	Computer Assisted Language Learning
CAUS	causative
Cl	clitic
CL	classifier
CMC	Computer-Mediated Communication
COMP	complement
Cop	copula
CP	complementiser phrase
c-structure	constituent structure
DAT	dative
DEF	definite
DF	discourse function
DU	dual
DOM	Differential Object Marking
ERG	ergative
ESL	English as a second language
FEM	feminine
FOC	focus
f-structure	functional structure
GEN	genitive
GEND	gender
GF	grammatical function
INF	infinitive
IP	inflectional phrase
JSL	Japanese as a second language
L1	first language
L2	second language

LFG	Lexical Functional Grammar
LOC	locative
MASC	masculine
NEG	negative
N	noun
NOM	nominative
NP	noun phrase
NUM	number
OBJ	object
OBJ θ	secondary object
OBL	oblique
OBL θ	oblique family
OVS	object–verb–subject (word order)
PALA	Processability Approaches to Language Acquisition
PASS	passive
PAUC	paucal
PERS	person
PL	plural
POL	polite
PP	preposition phrase
PROG	progressive
PT	Processability Theory
QP	question phrase
Q	question particle
QUE	question feature
QW	question word
S	sentence
SUBJ	subject
SG	singular
SLA	second language acquisition
SLI	Specific Language Impairment
SOV	subject–object–verb (word order)
SVO	subject–verb–object (word order)
TL	target language
TOP	topic
V	verb
VP	verb phrase
VSO	verb–subject–object (word order)
UG	Universal Grammar
XP	open phrase

PART I

Introducing and developing Processability Theory

Camilla Bettoni* and Bruno Di Biase**

*University of Verona, **University of Western Sydney

This initial part of the volume includes a presentation of PT's main tenets, which is designed first of all to highlight particular developments and proposals that have arisen since Pienemann (1998). Because much of the substantial progress in PT relies on developments in its two theoretical sources, in our exposition we give priority to Levelt's (1989) psycholinguistic Model and later developments (Bock & Levelt 1994; Levelt, Roelof & Meyer 1999; Levelt 2000) for language production, and to current LFG (Bresnan 2001; Dalrymple 2001; Dalrymple & Nikolaeva 2011; Falk 2001; Asudeh & Toivonen 2010) for language description, and allow them as much space as that reserved for the learner's developing path. Furthermore, this initial part is designed to give conceptual coherence and terminological consistency to the whole volume, thus avoiding repetitive introductions in each of the subsequent chapters. In this way, we hope to achieve two further aims: to illustrate PT's universal schedules of grammatical development and our 'prominence' proposal in a way that is consistent with its feeder disciplines, and also, crucially, to try to explain further the reasoning behind the schedules and the way they are connected, thus contributing to theory construction in SLA.

The progress in the two feeder disciplines that, in our view, bears most fruitful consequences for PT's developments and proposals concerns, on the one hand, (a) Levelt, Roelof & Meyer's (1999) Theory of Lexical Access, and (b) Bock & Levelt's (1994) specifications about the sequencing of the formulator's encoding procedures, with regard to language production; and on the other hand, (c) the formal representation of discourse functions (DFs from now on), and (d) the Lexical Mapping Theory in LFG's framework (Bresnan 2001; Dalrymple 2001; Dalrymple & Nikolaeva 2011; Falk 2001) with regard to linguistic knowledge. Consequently, in this part I, the main novelty in terms of contribution to theory construction derives from a coherent deployment of this progress in explaining the learner's developmental path by proposing our Prominence Hypothesis, as well as in attempting to solve incongruities deriving from Pienemann's continuing reliance on the older version of Levelt's Model, the earliest LFG, and derivational syntax (cf., Pienemann & Keßler 2011).

In relation to Levelt's Model, the Theory of Lexical Access has, crucially, introduced a third conceptual component to the original lemma-lexeme dichotomy. This development has its main application for PT in the Lexical Mapping Hypothesis by Pienemann, Di Biase & Kawaguchi (2005), whereby the (conceptual) lexical requirements of the verb contribute to driving the structural choices in the clause. Moreover, Bock & Levelt (1994) clearly show that, in the temporal course of language production, functional processing precedes positional processing. In PT terms this means that learners will first build up their syntactic frames and learn to assign grammatical functions (GFs from now on) to the retrieved lemmas, and only later will they relate morphological inflection to the constituents.

In LFG, the formalisation of the DFs, particularly Topic (TOP) and Focus (FOC) as syntacticised relations, has allowed for the development of a new dimension for PT, in so far as it can now formally represent promising areas – such as topicalisations, question formation, and other ways of attributing prominence – for investigating the learner's behaviour at the crucial intersection of syntax and discourse-driven choices. Furthermore, LFG's Lexical Mapping Theory contributes to explaining for PT how learners develop beyond rigid default mapping between thematic roles and GFs towards more flexible nondefault mapping in order to enhance expressivity and establish a different perspective or point of view on the event they intend to communicate. For instance, previous PT was silent on why learners who have acquired canonical word order may fail to produce passives despite their apparently pretty ordinary SV structure.

In presenting the learner's progress in developing their L2 grammar, we introduce several main innovations. The first is the separation of morphological development from syntactic development. This is different from Pienemann's own work (e.g., 1998, 2005, Pienemann & Kessler 2011) and much of PT work so far. The reason is that these two schedules appear to depend on two different sets of motivations. On the one hand, we have the original psycholinguistic procedures of Kempen & Hoenkamp (1987) assumed by Levelt (1989) and adopted by PT in Pienemann (1998), who shows how these processing procedures can be modelled in LFG by the mechanisms of feature unification. These procedures trace the developmental path of the learner's morphological marking beyond lexical learning over the hierarchical (phrasal, interphrasal and interclausal) levels of syntactic organisation (cf. ch. 1, § 4.1). On the other hand, the development of syntax depends on two different kinds of correspondences that formally relate three LFG parallel structures: argument structure, functional structure, and constituent structure (respectively a-structure, f-structure, and c-structure from now on). One set of these correspondences – that is, the mapping of c-structure elements (NP, VP, and so on) onto f-structure elements (SUBJ, OBJ, and so on) – describes the precedence relations of arguments according to the allocation of LFG's grammaticised DF, such as TOP and FOC (cf. ch. 1, § 4.2.1). The other set of correspondences

is guided by the principles of LFG's Lexical Mapping Theory, which accounts for the mapping of a-structure (a hierarchically organised set of semantic roles) to f-structure (a hierarchically organised set of GFs) (cf. ch. 1, § 4.2.2). Furthermore, by keeping morphological development separate from syntactic development, we are able to clarify and investigate issues involved in the interface between them (cf. ch. 1, § 4.3).

Other important innovations in chapter 1 concern the three hypotheses proposed in Pienemann, Di Biase & Kawaguchi's 2005 extension. In § 4.2, we abandon one of them, the Unmarked Alignment Hypothesis, and propose a reformulation of the other two, the Topic Hypothesis, and the Lexical Mapping Hypothesis. First, the Unmarked Alignment Hypothesis is abandoned because, at the initial grammatical stage, canonical word order and canonical (default) mapping between thematic roles and GFs do not necessarily entail each other. In fact, under pragmatically marked conditions a thematic role other than the highest in the hierarchy (e.g., the theme) may occupy the most prominent position in the string (i.e., the first). Secondly, in § 4.2.1, the original Topic Hypothesis is recast as Prominence Hypothesis, so that, parallel to the development of topicality in declarative sentences, it now explicitly includes the development of focality in interrogative sentences, thanks to the more general processing principle whereby any constituent can be made prominent by grammatical means. Among these means, together or separately, Levelt (1989) includes early appearance in the sentence (cf. 'linear precedence' in Sells (2001: 1), Choi's (2001) work on information structure, and Lee's (2001) work on word order), and mapping the role with the allocated prominence onto the highest-ranking GF (i.e., SUBJ) – as well as prosody, which however we do not treat because it clearly lays outside the scope of this volume. Thirdly, in § 4.2.2, our Lexical Mapping Hypothesis now includes an intermediate stage between the initial default mapping stage and the final nondefault mapping one. Thus our schedules for the development of syntax (based on the Prominence Hypothesis and the Lexical Mapping Hypothesis) now both hypothesise a middle stage which acts as a sort of lockpicker to open up their respective higher stage.

Finally, our presentation of PT in chapter 1 attempts to reflect more consistently a basic assumption about language development shared not only with many SLA researchers but also with L1 acquisition researchers and typologists. With Andersen (1984), Brown (1973), Keenan & Comrie (1977), Krashen (1982), and many others, PT assumes that the learner proceeds from least marked, feature-scant forms and structures towards more feature-rich, more specified and more marked forms and structures. Our contribution here, from a processability vantage point, attempts to throw some light on how learners proceed from obligatory defaults towards the skills needed to handle discourse-pragmatically induced and grammatically laden options in their L2 communication.

This path from obligatory grammar towards the deployment of a greater range of grammatical options is a characteristic that distinguishes the more advanced learners from beginner and intermediate learners, thus coming to terms with a perceived gap in PT, which has so far dealt more thoroughly with obligatory grammar in early interlanguage.

1

Processability Theory: theoretical bases and universal schedules

Camilla Bettoni* and Bruno Di Biase**

*University of Verona, **University of Western Sydney

1. Introduction

PT can be seen as a ‘progressive’ theory, that is, a theory capable of extending its domain, refining its concepts, making its key variables more operational, and attracting more research (Jordan 2004: 227). Even as we write, PT is expanding rapidly, and our presentation must inevitably be limited. But it is limited in four further ways. First, in this chapter we will mention PT’s history only when it helps explain some of the incongruities we try to eliminate, or justify our own choices. Secondly, our outline here is not intended as an independent introduction to the theory, in the sense that we will mention only minimally PT’s main scope, constructs and processes. These can be found in the original works by Pienemann (1998, 2005b; Pienemann, Di Biase & Kawaguchi 2005; including the more recent Pienemann & Kessler 2011, 2012) and in Pienemann’s own numerous shorter presentations (e.g., 2003, 2007) – although with regard to the latter a note should be added to the effect that they mostly refer to the acquisition of English, and rely on older versions of PT, barely touching upon the 2005 extension. Thirdly, although we will mention some problems in the theory, we do not intend to solve them all here. Nor, finally, can this chapter be read as a full review of the rich and varied PT literature.

On the positive side, our ambition here is to offer an outline of PT which is tightly anchored to its two updated psycholinguistic and theoretical linguistic bases and less dependent on the English language for exemplification.¹ So, we aim to offer a balanced synthesis: both critical in pointing out areas of weakness in PT, and enthusiastic in showing how its hypothesised universal schedule can generate paral-

1 Of course, we will continue to use English as the main L2 for illustrating PT, because it is the language familiar to most readers and most studied acquisitionally. Yet when we do so, care is taken to point out its typological peculiarities and present our discussion in such a way as to accommodate the widest possible crosslinguistic variation.

lel language-specific schedules widely across languages (as part II shows) and in a variety of situations (as indeed the chapters that follow in part III illustrate). Needless to say, our focus here is not so much on the details of PT's developmental schedules as on the reasoning behind them.

In sum, our first main focus here is on integration and coherence among what is at times separately and varyingly treated in PT literature: to wit

- between PT itself and its two theoretical bases (i.e., Levelt's psycholinguistic Model for language production, and LFG for language representation);
- between the original 1998 version by Pienemann and its 2005 extension by Pienemann, Di Biase & Kawaguchi; and
- among studies on different languages which are based on one or the other of these two versions.

This will help us offer our own proposals as a contribution to theory construction, which is the second main aim of this chapter. In turn, this will provide an opportunity for revisiting the earlier developmental schedules of three typologically representative languages in part II of the volume, and preparing the grounds for the new developments and explorations discussed in part III.

The remainder of this chapter is organised as follows: § 2 deals with language production (i.e., Levelt's Model, in § 2.1) and linguistic knowledge (i.e., Lexical Functional Grammar, in § 2.2) in relation to PT; § 3 summarises PT's key concepts; § 4 traces the learner's progress, first for morphology (§ 4.1), then syntax (§ 4.2) according to the Prominence Hypothesis (§ 4.2.1) and the Lexical Mapping Hypothesis (§ 4.2.2), and finally discusses how the morphological schedule and the two syntactic ones may interface; § 5 sets out some methodological issues for testing PT's hypotheses; and § 6 offers some concluding remarks.

2. Language production and linguistic knowledge

Closely paraphrasing Pienemann (1998, 2007), the underlying logic of PT is that L2 learners, at any stage of development, can produce only those L2 forms which the current state of their language processor can handle. This places PT among those SLA theories that see language primarily as a mental construct, like other generative approaches, but unlike most of them it considers the acquisition of an L2 as connected to real time human performance constraints such as those regarding speech processing. It is therefore crucial to base our understanding of language development on PT's two formal models, accounting for (a) language generation, that is, how the processor handles language, and (b) linguistic knowledge, that

is, what is language and how it may be represented in our mind/brain. Given that the anatomy and physiology of the human language processor are universal, if specific languages and their development are described according to a principled architecture, it may be possible to predict broadly similar grammatical trajectories for L2 development across languages.

For language generation, PT relies primarily on Levelt's Model (1989), a dynamic model accounting for language processing in real time and within human psychological constraints, such as the requirement for a very fast word access and, at the same time, the limitations of human memory. Thus a set of psycholinguistic universal constraints comes to bear on L2 acquisition, and provides a framework for PT's universal hierarchy of processing procedures, accounting for why learners seem to follow similar paths in their L2 grammatical development, across languages, whatever their L1 background. For linguistic knowledge, on the other hand, PT relies on LFG as originally conceived by Kaplan & Bresnan (1982) and further developed by Bresnan (2001), Dalrymple (2001), and Falk (2001) among many others over the last three decades. LFG provides a generative, explicit and well-defined, formal theory of language, which contributes towards solving for PT not only the problem of relying on a plausible mental representation of grammatical structure, but also Chomsky's well known 'logical problem': what is the origin of linguistic knowledge? Why do people end up knowing more than they can hear? As Pinker (2004: 949) puts it: "It is the question of how acquisition could work in principle – how a learner can correctly generalize from a finite sample of sentences in context to the infinite set of sentences that define the language from which the sample was drawn." For example, how does a learner know that words can be nouns, verbs, etc., or that in the sentence *The boy who loves Therèse is Indonesian*, Therèse may not be Indonesian despite the sequence *Therèse is Indonesian*? The interface between these two formal theories, Levelt's Model and LFG, allows PT to make language specific predictions about L2 development which can be tested empirically.

These two source theories of PT interface well because LFG is a constraint-based theory of generative grammar (Asudeh & Toivonen 2010) which intends to be psychologically plausible. Indeed, its suitability for psycholinguistic work is supported by the fact that LFG was chosen for language acquisition work, such as Pinker's (1984), as well as Levelt's Model of the speaker. There are good reasons for this compatibility. First, unlike other generativist frameworks, LFG's approach is decidedly lexicalist, declarative and nonderivational, which sits well with the psycholinguistic understanding that processing is formulated in one stage (Pickering, Branigan & McLean 2002), i.e., transformations are unlikely. Secondly, it explicitly connects competence with performance phenomena, in so far as, for example,

LFG's feature unification is related to performance in real time (cf. Rothman & VanPatten 2013). Thirdly, LFG's parallel projection architecture iconically suggests a representation of the dynamic processes temporally modelling language production. In particular, the fact that LFG's interest in typological questions of similarities and differences among languages has been central to its theoretical development (Asudeh & Toivonen 2010) finds parallels, *mutatis mutandis*, with PT's own history and current developments.

The next two sections (§§ 2.1-2.2) are meant to introduce the reader to those key features of PT's source disciplines that bear most directly on (i) its general architecture, and (ii) some of the issues mentioned in the volume. In no way do these two sections dispense the researcher interested in pursuing the finer details of PT from reading the original works by Levelt and Bresnan, and the ongoing updates by their respective teams – anymore than s/he is dispensed from reading Pienemann's own original works on PT.

2.1. Levelt's Model and PT

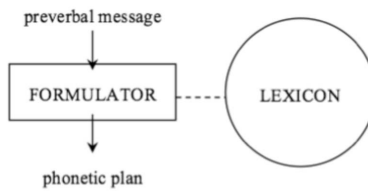
The debt to Levelt's Model is already fully acknowledged in the original PT version, when Pienemann (1998: §§ 2.4-2.5) highlights issues concerning the storage of grammatical information during language production, and the general psychological constraints that bear on language development. Thus we now focus on elements of language generation that help us understand PT's newer developments. These concern mainly two aspects of Levelt's Model: the lexicon and grammatical encoding. First, following developments in Levelt's Model itself (Levelt, Roelofs & Meyer 1999), a novel look at lexical access theory allows for a more precise characterisation of the lexicon and its central role in PT, which can now account for the acquisition of a wider range of features and constructions. If Pienemann's main concern in 1998 was the establishment of minimal requirements for reaching a stage, PT can now explore how learners proceed from the emergence of a structure to its mastery, or indeed from the emergence of one or two 'typical' structures in a stage to the mastery of the range of structures in that same stage. This contributes to accounting for development within a stage, which Mansouri & Håkansson (2007) began to explore as intrastage phenomena and we further characterise in this volume as progress through 'soft barriers' (cf. § 5 below, and further ch. 2, § 2 for English, and ch. 3, § 2.2 for Italian). That is, whereas Pienemann (1998) proposed modules other than PT for handling the complexities within a stage, we propose instead to integrate them into PT's explananda. Secondly, we wish to outline that part of Levelt's Model which bears more directly on PT's extension, because the extension of the theory by Pienemann, Di Biase & Kawaguchi's (2005) dealt more fully with its LFG formalism rather than discuss the details of language production.

Levelt's Model, along with other speech production models that have been proposed from various theoretical viewpoints over the last 25 years or so, share the understanding that the speech production process can be described as consisting of three broad components: conceptualisation, formulation and articulation (cf. Grosjean & Li 2013: 51). Levelt assumes that when we intend to say something we select in the conceptualiser the information whose expression may realise our communicative goals – the conceptualiser being the processor where the preverbal message is generated and then fed to the formulator, as the fragments of the preverbal message become available. Since any state of affairs can be expressed in many different ways, in the conceptualiser we also plan the form of the message, in the sense that here we select not only the language and register but also the appropriate speech acts or rhetorical device required (assertion, question and so on), we tag topic and focus, mark the referents as given or new, and so on. Thus, the conceptualiser's output (i.e., the preverbal message) already includes information on a number of choices, including the relative prominence of its elements. All this presents no problem for adult L2 learners, who are already fully competent speakers of their L1, because, in order to produce a preverbal message they can rely on the same conceptualiser for either language, as De Bot (1992) maintains in adapting Levelt's Model of language production to bilingual speakers.² On the other hand, grammatical learning begins when then the formulator – which is language-specific – receives input from the conceptualiser, and has the task of mapping the preverbal message onto the appropriate linguistic form (i.e., the formulator is responsible for grammatical encoding), and preparing (through phonological encoding) the phonetic plan, as represented in (1). This task is performed by retrieving out of the lexicon the stored entries that best fulfil the conditions required by the preverbal message. Thus, just as the learner has to learn, however gradually, a language-specific lexicon, so do they also have to build, contextually, a language-specific formulator to process it. Levelt (1989: 103) assumes, in fact, that “[t]here are different formulators for different languages”.

Let us then take a look in turn at how lexical entries are stored in the lexicon, and then processed in the formulator.

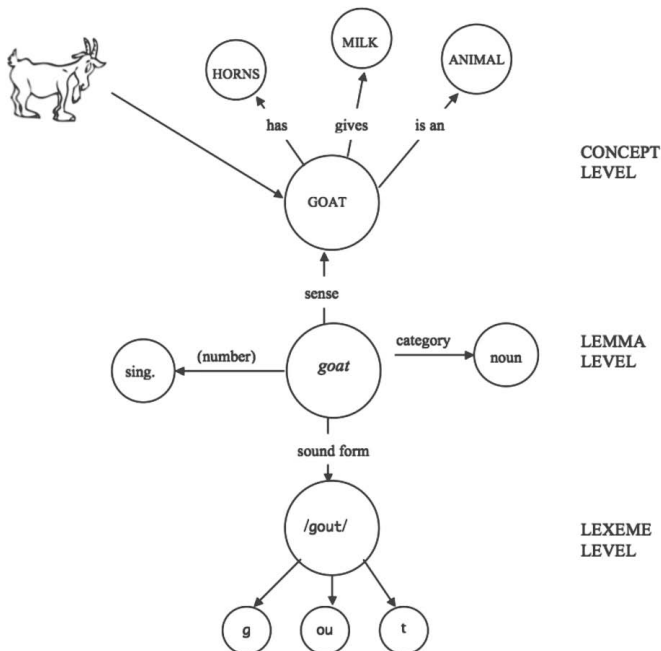
2 Once the language has been specified in the conceptualiser, the output is language-specific because “[w]hatever the speaker tends to express, it should ultimately be cast in terms of lexical concepts, that is, concepts for which there exist words in the target language” (Levelt, Roeleof & Meyer 1999: 8). Despite considerable growth in psycholinguistic experimental research of spoken language processing in bilinguals, “a comprehensive model remains to be developed” (Grosjean & Li 2013: 52).

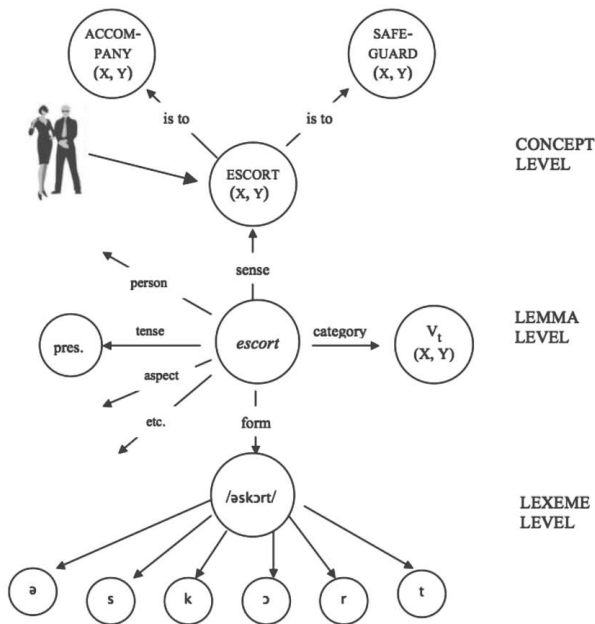
(1) *Language production from preverbal message to phonetic plan (after Levelt 1989: 9)*



In the lexicon of the adult L1 speaker, according to Levelt, Roelofs & Meyer's (1999) Theory of Lexical Access, words are stored with the full bundle of features involving three types of information, distributed in a three-level system: the conceptual level, the lemma level and the lexeme level. We should point out that the conceptual level was not present in Pienemann's 1998 application of Levelt's Model to PT, nor has it been incorporated in Pienemann's subsequent versions, which retain Levelt's (1989) two-tier system of the lemma and the lexeme. In Levelt, Roelofs & Meyer's (1999) three-tier system, instead, conceptual information include elements associated, for example, with the semantic roles of V (e.g., <eater, eaten> for *eat*), rather than infor-

(2) *A fraction of the lexical network for the word goat (after Bock & Levelt 1994: 951)*



(3) A fraction of the lexical network for the word *escort* (after Levelt, Roelofs & Meyer 1999: 4)

mation about the number and type of arguments only. Notably, PT's extension would not work without the conceptual tier of the lemma.

In (2)-(3) we give a simplified representation for two lexical entries, the N *goat* and the V *escort*, following Bock & Levelt (1994: 950-952), and Levelt, Roelofs & Meyer (1999: 3-4).

First, at the conceptual level, knowing a word involves knowing its meaning. About a *goat* we know it is a kind of domestic animal that produces milk, etc., and also that it typically de-selects certain other words such as *think* or *smile* typically reserved for humans, etc. About *escort* we know that it is an action related to accompanying, guiding, etc., that it requires the two semantic roles of agent and patient, etc. These are properties of our concepts GOAT and ESCORT.

Secondly, at the lemma level, a word has syntactic properties, a bundle of grammatical features – including also combinatorial information (cf. Sells 2001; Kim & Sells 2008) – which place it in its syntactic frame. The English word *goat* is an N. Its Italian equivalent *capra* is also an N, but in addition it has feminine syntactic gender. The word *escort* is here a V, and Vs are specified for (in other words, they select) the arguments they require, corresponding to their semantic roles (<agent, patient> in this case); thus about *escort* we know that it typically

requires a SUBJ and an OBJ on which to map the required semantic roles. As will be clear further on, these lexically determined requirements of the V for argument and function specifications are of crucial importance for understanding, and operationalising, PT's Lexical Mapping Hypothesis, first advanced by Pienemann, Di Biase & Kawaguchi's 2005 (cf. § 4.2.2 below). Moreover, many lemmas require so-called diacritic features that have to be learned. For example, in English V lemmas have diacritic features of person, number, tense, and mood, which must be valued for further encoding. Hence the lemma *escort* will be realised phonologically as *escort*, *escorts*, *escorted*, or *escorting* depending on the values of its diacritic features. Some values of these features derive from conceptual representation, as when English Vs are marked for tense, or Ns for number, others may be set in the course of grammatical encoding, as we will shortly see below.

Thirdly, at the lexeme level, knowing a word means knowing its formal properties that is, its morphological and phonological shape. The word *goat* is monomorphemic and consists of three phonological segments: /g/, /ou/, and /t/, whereas the Italian word *capra* consists of two morphemes, a stem (*capr-*) and a suffix (-a), and five phonological segments: /k/, /a/, /p/, /r/, and /a/. Likewise, in (2), the nodes at the form level represent phonemic segments.

In Levelt's Model, it is the lexicon – with its associated semantic, grammatical and phonological information – that primes the procedures and feeds forward the encoders. As we move on to grammatical encoding, let us remember that all this information is characteristically stored in the mature native speaker's mental lexicon, but learners build up their L2 lexicon gradually. If, on acquiring a new word, learners may quickly be able to associate a 'meaning' at the conceptual level with some (phonological) 'form' at the lexeme level, the same is certainly not true for the lemma level, where features and values may take a long time to emerge³ and even longer to master. So, whereas it will not take long for English learners of Italian, for example, to learn the word *capra* and get to know what it means (*goat*), it will take them longer to learn that the N *capra* has the value 'feminine' with respect to the N feature 'gender', which in turns combines with 'number' information, and hence, if they want to refer to more than one, they need to use the form *capre* (not *capras*).

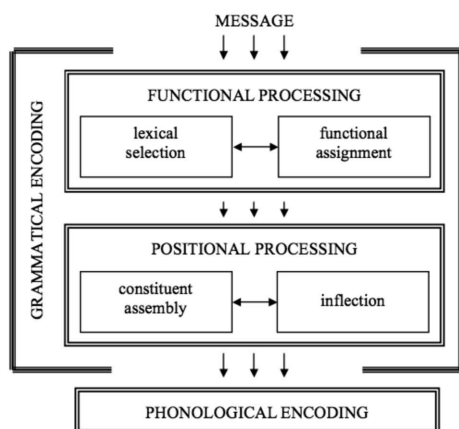
So, unlike many of the components at the conceptual level, which may find a surprisingly high degree of commonality across languages, diacritic features and their values at the lemma level tend to pop up in peculiarly idiosyncratic mixes and with surprisingly different exponents requiring unification or merger over varying degrees of syntactic distance – all very much in language-specific ways. These idiosyncrasies of the L2 lexical features fuel one of the least understood and

3 Emergence can be understood here both as annotation of each feature-value pair in the lexicon and their retrieval during production.

perhaps most confusing issues in PT, as well as in other approaches. The question is why is it that structures that are apparently at the same stage of development (e.g., marking some sort of aspectual feature with *-ing* or tense feature with *-ed*, where both are category level morphemes) should emerge so far apart from each other? Within PT this has been variously cast as ‘intrastages’ or ‘steps within a stage’. In our treatment we choose the term ‘soft barriers’ to indicate that the learner, having overcome the ‘hard barriers’ and learned the skill to go past the important constraints imposed by a stage, will then learn more fine-grained lexically-based distinctions. Similarly at the lexeme level, there may well be peculiar lexical class distinctions interacting with specific features at the lemma level that may be rather difficult to acquire for adult learners (cf. Di Biase 2008 for one example of such lexical class-based distinctions, affecting number in Italian nouns and adjectives). We will see below in § 2.2 how the lexicalist orientation of LFG, which indicates feature-value pairs formally, is particularly apt in providing an efficient representation (presence or absence of a particular feature/value) of such phenomena.⁴ And we will come across specific examples of soft barriers in the language-specific chapters in Part II.

The formulator – to resume our presentation of Levelt’s Model – encodes the utterance first grammatically and then phonologically, as shown in (4). We are interested here in grammatical encoding, whose processes create the skeleton of the

(4) *Components of grammatical processing (after Bock & Levelt 1994: 946)*



⁴ As well as exploited by our proposal in the context of L2 learners’ progress, LFG’s lexicalist orientation, efficiently represented in formal annotation, is pointed out by Schwarze (2002) in the context of (historical) language change.

utterance. These processes are grouped into two sets – one functional and the other positional – each set with its own subcomponents. It may be worth noting that in (4) the arrows represent the flow of information during production (and, in reverse order, comprehension) and depict activation trajectories, whereas in (2)–(3) they represent types of connections within the network and depict permanent relationships in a store.

Functional processing has two subcomponents: lexical selection and functional assignment. Given a lexical concept to be expressed, lexical selection involves retrieving a word, or more specifically a lemma, from the lexicon. Functional assignment involves creating the appropriate syntactic environment for the words by assigning them their syntactic functions. For example, upon selecting the lemma *escort*, its syntax – it is a transitive V with two argument positions, corresponding to its two semantic arguments – will become available for further grammatical encoding (Levelt, Roelofs & Meyer 1999: 4): which of the two arguments will serve as SUBJ, which as OBJ?

Functional assignment is controlled by two kinds of information represented in the message. First, the eventuality conceived in the conceptualiser is associated with thematic or event roles, such as agent (the instigator of an event), patient or theme (the person or object that is affected or moved). This explains why, in organising an utterance, the V lemma chosen during lexical selection is central over other lemmas. Second, the relative prominence among the participants in the event is associated with discourse or attentional roles. These organise the informational distribution in the utterance so as to direct the listener's attention to its components. As Bock & Levelt (1994: 964–965) comment, there are “seductive correspondences” between both thematic and discourse roles and grammatical functions (GFs). That is, agent is most often SUBJ, beneficiary is OBJ, etc., as in (5a), although violations to these default correspondences are possible, as in (5b) or (5c), depending on discourse-pragmatic information tagged in the preverbal message.

- (5) a. Romeo gives a rose to Juliet
- b. Juliet is given a rose by Romeo
- c. Juliet receives a rose from Romeo

Likewise, elements expressing given (or topical) information, which are more readily available, often appear early in the sentence and have great affinity with SUBJ.⁵ This is shown in (6), where the same propositional content is expressed with diffe-

5 The developments in LFG since Bresnan (2001) go a long way in representing these processing computation issues by having incorporated hierarchical semantic roles in Lexical Mapping Theory and syntacticised discourse functions in its *f*-structure, as discussed in § 2.2 below.

rent prominence in (6a) and (6b) by assigning the SUBJ function to a different topical element:

- | | | |
|-----|-------------------------------------|--------------------------------------|
| (6) | a. [what's going on with the dog?] | the dog is chasing the cats |
| | b. [what's going on with the cats?] | the cats are being chased by the dog |

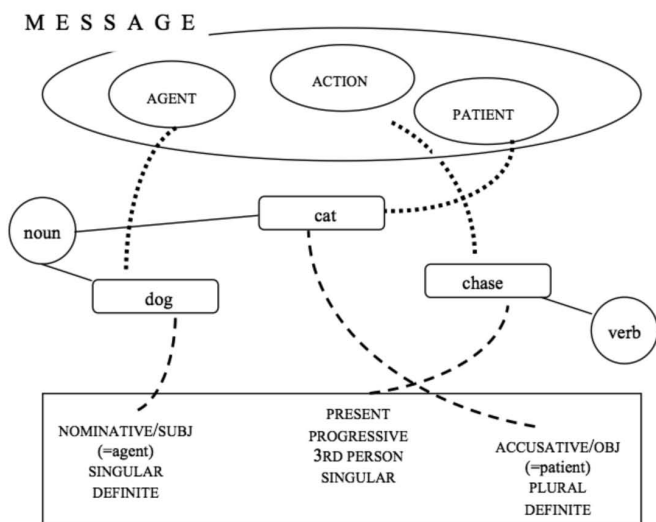
For the number and type of functions assigned during functional assignment, we refer to their LFG representation reported in § 2.2 and shown in (21). Suffice it to say here that, although functions are universal, they may be marked differently in different languages: morphologically by case markers, clitic pronouns, etc., or structurally by position, as shown in (23)-(24). These different ways are not mutually exclusive, and indeed many languages use a combination of means. Even a highly configurational language like English marks the case of personal pronouns morphologically (e.g., by distinguishing between *I* and *me*, *she* and *her*).

Finally, during functional processing, the combination of lexical selection and functional assignment also specifies the value requirements for the diacritic features of individual lemmas. For example, if the speaker intends to produce the sentence in (6a) above, upon selecting the V *chase* and the Ns *dog* and *cat* for expressing this current eventuality in the present involving *dog* as agent and *cat* as patient, functional assignment will determine not only the grammatical relations between the lemmas (i.e., *dog* (the chaser) is SUBJ and *cat* (the chased) is OBJ of *chase*), but also the values of the diacritic features (i.e., *dog*, referring to a single referent, is realised in its singular form as *dog*; *cat*, referring to more than one referent, is realised in its plural form as *cats*; and *chase* is realised in its present, progressive form *chasing* with the auxiliary carrying the required SUBJ number information with singular value).

In sum, functional processing yields an available and activated set of lemmas and a set of abstract syntactic functions, which are linked together via the argument structure of the lemmas, notably that of V (cf. (7) for a schematic illustration of the product of functional processing of the sentence *the dog is chasing the cats*). All this material (i.e., abstract relations or linkages among elements) may contain some indication of the relative prominence assigned to various components, but it is not ordered in any sequence, except for the order in which the fragments become available from the conceptualizer. To convert into an utterance, the fragments of this partial, incomplete structure do not go into the phonological encoder directly as they come out of functional processing, but they are stored temporarily in the memory buffer. The product of functional processing must now be processed positionally.

Positional processing, like functional processing, also has two subcomponents: constituent assembly and inflection. Both involve the creation of a set of

(7) *The product of functional processing for the dog is chasing the cats (after Bock & Levelt 1994: 968)*



slots which are ordered: the former for lexemes, the latter for morphemes (Bock & Levelt 1994).

Constituent assembly fixes the linear order of word production and captures dependencies among syntactic functions. Ordering is necessary because the output of functional assignment carries no intrinsic order.⁶ This becomes clearest not with English, a highly configurational language which marks grammatical functions (GFs) by position, but with less configurational languages, whose constituents may appear in different positions serving the same GFs, often signalled by differences in case marking. Russian and Latin are such languages. In (8) for example, by marking the SUBJ (*Paul*) of the V *amo* ('love') as NOM case by means of the *-us* morpheme, and the OBJ (*Mary*) as ACC case by means of the morpheme *-am*, Latin can place either anywhere, as required by discourse or pragmatics, without changing the propositional content of the message:

- (8) a. *Paulus Mariam amat*
 b. *Mariam Paulus amat*

⁶ Also LFG's *f*-structure is not intrinsically ordered (cf. § 2.2).

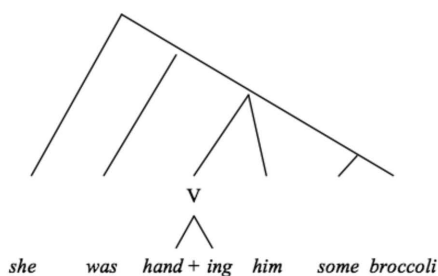
Establishing dependencies among words means organising phrase groupings in a hierarchy. Without them, as Bock & Levelt (1994: 969) point out, there would be, for example, no means to segment sentences such as (9) appropriately, where the listener knows that it is not to be understood that it is the boy who blushed, despite the linear sequence showing *the boy blushed*:

(9) the girl who kissed the boy blushed

This hierarchical organisation (namely, the phrase structure) is assembled bit by bit under the control of the syntactic functions and the grammatical categories of the lemmas that realise them. This means that, for example, given the nominative function and a N lemma to fill it, adequate information is available to create a SUBJ NP in the appropriate position in an utterance. As we will see further on, Bock & Levelt's claim that functional assignment precedes positional constituent assembly in the temporal sequence of language production by the native speaker is crucial for understanding learners' development beyond canonical word order.

Inflection is the last grammatical encoding process, and involves the generation of fine-grained details at the lowest level of the hierarchy of phrasal constituents, as shown in (10). This is a thorny issue, not yet solved in all its facets in Levelt's Model. The debate is around two questions: first, whether in such cases as the English *handing* or Italian *capra* the mental lexicon stores the whole word or its morpheme components; and secondly, whether to consider under inflection not only inflection proper, but also the formulation of function words often associated with grammatical phrases such as determiners for NPs, auxiliaries for VPs, and prepositions for PPs (Bock & Levelt 1994: 972). Suffice it to say here that LFG's principle of lexical integrity considers words as atoms from the point of view of syntax, that is, they are no further divisible into smaller syntactic units (cf. § 2.2). As stated by Asudeh & Toivonen (2010: 430), "[t]he syntax is [...] blind to the internal structure of words and it sees only their category". Furthermore, as mentioned above, certain lemmas carry specifications about diacritic features to be valued inflectionally. In some cases these specifications may be under the control of conceptual elements, as when Vs are specified for tense. In other cases the control is syntactic, as when there are dependencies among inflectional features. So, in the sentence in (10) speakers say *she **was** handing him some broccoli*, rather than *she **were***, because here two constituents of the sentence reflect a value (i.e., third singular) of some feature (i.e., person) that triggers inflectional variation. These constituents need not be adjacent. What is necessary is that the agreeing constituents stand in appropriate syntactic-functional relationships. In this English example, the agreement operates between the head of the SUBJ NP and the finite V.

(10) *Constituent hierarchy of she was handing him some broccoli (after Bock & Levelt 1994: 946)*



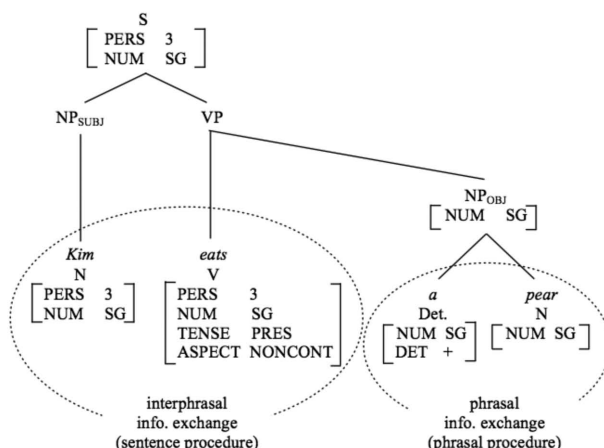
In sum, we can say that, first, functional processing integrates a set of lemma specifications with a set of syntactic functions. Its output is a set of abstract relations and properties which guides the creation of a frame for positioning words. Subsequently, positional processing places words and their inflections into that frame. The output of functional processing, then, is an ordered set of lexemes, formally realising the abstract relations of the functional specifications.

Why do we need to comprehend this complex process in order to understand the way PT explains the learner's progress in acquiring an L2? Though the reason may become clearer in §§ 3-4, we can anticipate here that, whereas mature L1 speakers are able to activate all the encoder's components effortlessly, L2 learners must build them up gradually. If De Bot (1992) is correct in saying that bilingual speakers operate with a different formulator for each language, while the L2 formulator is under construction learners will be able to produce only those structures that can be processed by the components already in place – as well as the lexical material already stored.

Let us now go back to the native speaker. In Levelt's Model this complex process of grammatical encoding for language generation assumes that grammatical information activated by one procedure needs to be stored temporarily in a specialised memory buffer in order to be used by another procedure, so that the two lots of activated information can then be compared by yet another procedure that builds the output of the first two procedures together. How does this process unfold? Following the Incremental Procedural Grammar developed by Kempen & Hoenkamp (1987), Levelt (1989) proposes that grammatical encoding in mature monolingual speakers follows the sequence in (11):

- (11) a. the lemma
- b. the category procedure
- c. the phrasal procedure
- d. the sentence procedure

(12) An illustration of processing hierarchy for *Kim eats a pear*: phrasal and interphrasal procedures



Thus, upon selecting the lemma, the category procedure is instigated, assigning a lexical category to the lemma. Then the category of the head lemma will instigate a phrasal procedure, resulting in a phrase. With the activation of the sentence procedure, phrases in turn acquire their function according to the syntactic frame of their head lemmas. To illustrate these steps with an example, in *Kim eats a pear*, shown in (12), first the lemma *Kim* needs to be assigned to the lexical category N, and its diacritic features (number and person) returned with their respective values (singular and third person). Then the lemma *eats* needs to be assigned to the lexical category V, and its diacritic features (number, person, tense, and aspect) annotated with their respective values (singular, third, present, and noncontinuous). Further, in order to achieve the agreement between the two phrases *Kim* and *eats*, information must be exchanged between them, and the values of the features they share (number and person) must be compatible. Likewise, in generating the NP *a pear*, the selection of the lemma *a* partly⁷ depends on the value (singular) of the diacritic feature of the phrase head lemma *pear*, because the common values they share must be checked against each other, at the phrasal node, for compatibility.⁸ In this case the value of the diacritic feature of *pear* is stored by the category pro-

7 The choice of *a* vs *the* also depends on the information tagged in the 'discourse model' (in the conceptualiser) determining its new vs old, definite vs indefinite values, and so on.

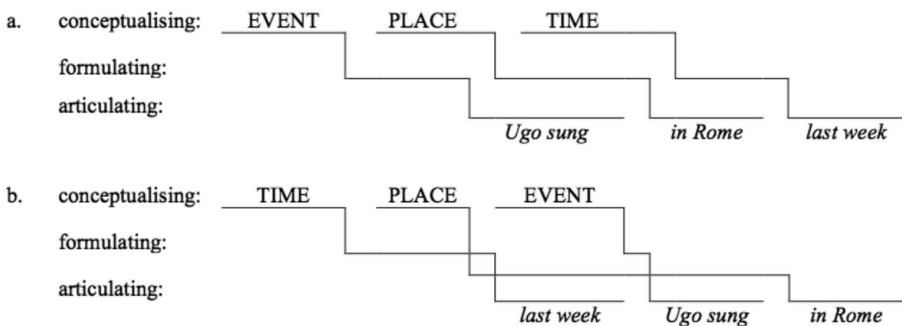
8 This notational innovation showing only the unifying features and their values at the node where unification is assumed to occur was introduced in LFG by Andrews & Manning (1999).

cedure until it is checked against that of the modifier *a*. Finally, in order to build up the sentence, a GF must be assigned to the two newly created NPs, that is SUBJ for *Kim* and OBJ for *a pear*. This matching, or exchange, of information regarding the values of shared diacritic features among the elements of a sentence is called ‘feature unification’ in LFG terminology (cf. § 2.2). It is a key LFG concept used in Levelt’s Model of the (monolingual) speaker and incorporated into PT to account for the gradual process of L2 development.

Kempen & Hoenkamp’s (1987) Incremental Procedural Grammar also assumes that the whole four-step sequence from activating the simple lemma to the sentence procedure is implicational. This means, for example, that in order to activate the phrasal procedure, both the lemma and the category procedures must be activated, but that the sentence procedure need not be active as yet.

Furthermore, the whole process of language generation is incremental. This means that all processors can operate simultaneously in parallel, but they all work independently on different language fragments of the utterance under construction. The final order of articulation may follow the sequence in which fragments are conceptualised, as in (13a), or it may not, as in (13b). Although here the speaker has decided to make some specific circumstance (in this case the context of the event) more prominent, the English formulator will not allow the fronting of both the time fragment and the place fragment (which would produce **last week in Rome Ugo sang*). So the formulator, being the specialist processing component charged with generating grammatically acceptable sentences, will order the fragments accordingly, and produce *last week Ugo sang in Rome*. These types of asynchronies do not create a particular problem or cost for the native speaker, who has automatised the necessary procedures to handle diverse ways of organising the mes-

(13) Incremental production (a) without and (b) with inversion of order (after Kempen & Hoenkamp 1987, as cited in Levelt 1989: 25)



sage online and delivering it in real time.⁹ But the processing cost of this very fast process, averaging between two and three words per second (Levelt 1989: 22), can be very high for the learner.

Finally, the theoretical assumption of incremental processing (i.e., of parallel processing activities in the different components of speech generation) hinges on automaticity. Automatic processes have great advantages: they are executed without conscious awareness; they are quick; and they run on their own specialised resources, which means that they do not share resources, so they can run in parallel without mutual interference. If each processor were to require access to attentional resources such as Working Memory, speaking would not be fluent, and between one articulated fragment and the next there would be silences devoted to processing. On the other hand controlled processing demands attentional resources, because we can only attend to a few things at a time. Attending to the process requires a certain level of awareness of what we are doing. Thus human controlled processing tends to be serial rather than incremental in nature, and is therefore slower. Its advantage is that it is not entirely fixated in memory; in fact, controlled processing is quite flexible and adaptable to the requirements of the task at hand. Levelt (1989) accounts for the amazing speed of the speech production process by assuming it is largely automatic. Only the message generation and monitoring activities are controlled processes, that is, normally requiring the speaker's ongoing attention, whereas grammatical and phonological encoding, articulation, and lexical access are usually automatic – the exceptions being, for instance, the retrieval of infrequent words (cf. Poulisse 1997). Here is where the difference between the native speaker and the L2 learner is crucial. Whereas the native speaker effortlessly generates fluent speech, the learner proceeds gradually from painfully slow retrieval of lexical items towards ever more complete grammatical encoding in an increasingly automatic way. The relevance of automaticity versus (attended) control in speech production is a fruitful area for future research.

This is then, in its basic outline, Levelt's Model for language generation in the mature native speaker. What about L2 learners? Summing up, as they develop their interlanguage, learners need to

- (14) a. build up the lexical store, and include in it fully mature lexical items; that is, not only increasing the number of words with their meanings and sound forms, but also storing richer lemmas with their relevant diacritic features (semantic, grammatical and formal, as well as categorical and combinatorial) and specific values;

9 Even for native speakers processing less preferred options involves more time, however minimally. E.g., Weyerts et al. (2002) report that German native speakers prefer SVO orders to SOV orders and account for this in terms of the relative higher processing cost of latter as against the former.

- b. learn to retrieve and encode these lemmas functionally (LFG's *f*-structure) and positionally (LFG's *c*-structure) (cf. § 2.2 below); and
- c. automatise increasing numbers of retrieval and encoding processing components for fluent speaking, so as to free up greater attentional resources for semantic and discourse-pragmatic processing.

The sequence in which learners gradually learn to enrich their lexical storage, activate more grammatical procedures, bottom up, and automatise them is precisely PT's domain (cf. § 3). In 1998 Pienemann's main concern was with the learner's progressive activation of grammatical procedures in the formulator. This allowed for PT's explanation of the development of obligatory morphological and morpho-syntactic structures. Then in 2005 Pienemann, Di Biase & Kawaguchi began to focus more sharply on the role played by (a) the conceptualiser in preparing its output with a variety of discourse-pragmatic perspectives, and (b) the lexicon in the grammaticalisation process. Bringing into PT this element of choice the speaker may have in selecting alternative discourse-pragmatic perspectives and alternative lexical items allows for new explanations regarding optional syntactic structures. In this volume the structural consequences of the speaker's choices will be taken a step further (cf. our own formulation of the Prominence Hypothesis and the Lexical Mapping Hypothesis in § 4.2.1 and § 4.2.2 respectively). With regard to automatisation, very little work has so far been done apart from Pienemann (2002) and Kawaguchi & Di Biase (2012). This is an important line of research worth pursuing within PT.

But before describing PT's developmental stages based on implicational procedural skills identified in Levelt's Model, we need to introduce some basic notions of LFG, and integrate them – as far as possible – within this model.

2.2. *Lexical Functional Grammar and PT*

LFG¹⁰ is committed to the interface between linguistic knowledge and language processing, and is therefore designed to account for linguistic knowledge in a way that is compatible with the architecture of the language processor (Kaplan & Bresnan 1982: 177). This must be the sort of commitment that prompted Levelt to use Bresnan's lexically-based grammatical theory for his Model's lexicalist approach – in the sense that, as we have seen, Levelt assumes that lexical selection activates syntax, that is, it drives the procedures for building the syntactic frame for the current utterance (cf. Kormos 2006: 9–10). So, in the speaker's mental lexicon each

10 Cf. Asudeh & Toivonen (2010) for a concise, up-to-date presentation of LFG.

lexical entry is associated not only with meanings and forms, but also with the full set of their syntactic information.

In addition to its lexicalist assumption, another important principle of LFG is its clear distinction between levels of linguistic representation that a formal model requires in order to describe the complex architecture of natural language adequately. Since the sentence is an expression of several types of linguistic information (semantic, pragmatic, syntactic, phonetic/phonological, etc.), there can be several theoretically distinct structures of formal representation: semantic structure, information structure, and phonological structure, as well as the three syntactic levels: argument structure, functional structure and constituent structure. All these levels exist simultaneously and in parallel,¹¹ each with their own distinct grammatical module, with characteristic primitives and formal representation. LFG's architecture "postulates a number of simple data structures with mappings defining the relationship between the structures [...]" (it is thus typically referred to as a *Parallel Projection Architecture*" (Asudeh & Toivonen 2010: 426, original italics). We will concentrate here on the three syntactic structures: a-structure, f-structure, and c-structure. LFG work on the other levels is more recent and not yet applied to PT, so it will not be further mentioned here except for some rudimentary information structure, which is important for our Prominence Hypothesis (cf. § 4.2.1 below).¹² We first consider the syntactic structures separately, with their specific properties that make each level of representation different and independent from each other, and then we show how they are linked, or mapped, onto one another.

Beginning with a-structure, this level of representation determines many of the basic properties of the sentence in which a predicate occurs – the predicate being the word, typically a V, which names the action, event or state described by the sentence. The a-structure of a predicate encodes information about the number and type of arguments selected by that predicate, as shown in (15). Arguments are thus assigned lexically through the meaning of the V.

- (15) *run* <agent>
 eat <agent, patient>
 love <experiencer, stimulus>
 give <agent, theme, recipient>

11 A good discussion and motivation for this architecture may be found in Falk (2001).

12 Further LFG work in this direction might help PT's development, especially concerning i-structure. Cf., e.g., the debt to Choi (1999, 2001) for the Prominence Hypothesis, and to Mycock (2007) and Dalrymple & Nikolaeva (2011) for chh. 7 and 8, this volume.

Because argumenthood is a semantic concept, it is not always as easy to determine the arguments of a predicate as one might expect, and many label sets have been proposed. Suffice it to say here that LFG generally follows Jackendoff (1972), and others, in defining a hierarchy among them which is formally based on two very broad aspects of the way we conceptualise the meaning of V_s: one based on action, the other on space. In the action conceptualisation, an agent has primacy over a beneficiary, because when both are present, the beneficiary is affected by something the agent does. In the space conceptualisation, the instigator has primacy over the theme, which in turn is more prominent than the path, etc. By stipulating also that the action conceptualisation takes over the spatial one, the thematic hierarchy in (16) is derived, typologically validated by Keenan & Comrie (1977) and Hopper & Thompson (1980), among others, and used by Bresnan (2001: 307):¹³

(16) agent > beneficiary > experiencer/goal > instrument > patient/theme > locative

Moving on to *f*-structure, this level of linguistic representation includes, for every sentence, all the grammatical information needed to interpret the sentence semantically. It consists of two types of information about the syntactic elements (i.e., words and phrases) of a sentence:

- information about grammatical relationships between them; and
- information about their grammatical properties (i.e., features).

That is, in *f*-structure, abstract GFs and diacritic features attempt to capture universal syntactic principles that may vary crosslinguistically at other levels of representation. We will now have a closer look at these two types of information.

In *f*-structure encoding, let us look first at GFs, which LFG, uniquely among grammatical theories, considers as primitives of the theory.¹⁴ GFs are cross-classified in several ways. The most basic ones are the argument functions, which express the arguments of predicates. Indeed, they may be directly selected, governed, by the predicate. They are: SUBJ (subject), OBJ (object), OBJθ(secondary object), the OBLθ (oblique) family of functions, COMP and XCOMP (complement). Among them a fundamental distinction is made between core functions, which are SUBJ,

¹³ Because a thematic hierarchy is not a primitive construct, other rankings have been proposed – all of which, however, almost without exception, rank agent highest. For a discussion of criteria used in constructing a thematic hierarchy and ranking semantic roles, cf. Levin & Rappaport Hovav (2005: ch. 6).

¹⁴ For basic definition and examples of LFG's GFs, cf. Asudeh & Toivonen (2010: 432).

OBJ and OBJ θ , and noncore functions, which are OBL θ , COMP and XCOMP (cf. also (21) below). Core functions are associated with the central participants of the eventuality expressed by V, and are usually distinguished formally from noncore functions. In English, for example, core arguments have canonical c-structure positions which are occupied only by NPs and DPs; noncore functions are generally expressed by other c-structure categories, such as OBLs by PPs. For example, in the sentence in (17a), *Consuelo* and *her sentiments*, respectively SUBJ and OBJ, are core GFs subcategorised by V, associated with NPs in c-structure, and obligatorily present in the sentence. On the other hand, *to Pablo*, an OBL θ , is a noncore GF associated with PP in c-structure, and optional, as can be seen from the fact that in (17b) it can be left out without affecting grammaticality.

- (17) a. Consuelo expressed her sentiments to Pablo quite nicely in writing
 b. Consuelo expressed her sentiments quite nicely in writing

Argument functions, like the thematic roles of a-structure, are also hierarchical: SUBJ > OBJ > OBJ θ , etc. (Keenan & Comrie 1977; Bresnan 2001). The hierarchical organisation of these two levels of structures (i.e., a-structure and f-structure) is relevant from a processing point of view, as we shall see later.

All GFs other than argument functions are not specifically selected by the predicate but occur freely, subject to other constraints of the theory (Asudeh & Toivonen 2010: 431). One such nonargument function is ADJ (adjunct). Whereas argument functions allow only single instances (e.g., there can be only one SUBJ per sentence), ADJ allows multiple instances. For example, in the sentences in (17) there are two ADJs (*quite nicely* and *in writing*).

All GFs mentioned so far, whether argument or nonargument ones, represent the clause-internal aspect of syntactic organisation. However, GFs can also relate to the wider discourse. So, as a secondary function, a syntactic element can also relate to the role its clause may have in the wider discourse structure. These secondary syntactic functions are discourse functions (DFs) or overlay functions. They are three: TOP (topic), relating the topic of the discourse, mainly old or shared information; FOC (focus), expressing new information; and SUBJ, which is the only DF that, being also an argument function, is also a governable GF (i.e., selected directly by the predicate). SUBJ is often identified in many languages as the default discourse topic.¹⁵ The TOP and FOC functions are not selected by the predicate and map indirectly to f-structure in the sense that they must be co-referential with, or anaphorically linked to, a nondiscourse GF. For example, in (17a-b) *Consuelo* is SUBJ (and default TOP), whereas in (18) the same propositional content topica-

15 For a discussion of subjecthood within the LFG framework, cf. Falk (2006).

lises *in writing*, which is TOP and ADJ; and in Bresnan's (2001: 97) example in (19b), the preposed NP *Rosie* is both FOC and OBJ of its sentence, with FOC relating this sentence to a previous element in (19a):

(18) *in writing* Consuelo expressed her sentiments quite nicely

- (19) a. what did you name your cat?
b. Rosie I named her

Note that DFs are not part of discourse representation, any more than nonDFs are part of lexical semantics, or any more than the notions of SUBJ and OBJ, or 'nominative' and 'accusative', need references outside the sentence in which they operate. DFs are somewhat similar to pronouns, for instance, in so far as they may have an argumental function in the sentence, but refer anaphorically to an entity or referent mentioned previously. It may be worth reiterating here that, in the LFG framework, DFs are *overlay* functions strictly *within* the sentence, and that by default (i.e., in the absence of other GFs specifically marked as TOP), TOP is overlaid to SUBJ, which is both a GF and a DF. These notions may become clearer in the context of Levelt's Model. As we have already mentioned in § 2.1, there are seductive correspondences between thematic and discourse roles on the one hand, and GFs on the other. That is, agents tend to be SUBJs, and elements expressing topical information tend to appear early in the sentence and have great affinity with SUBJ, a function that allows them to lead in the utterance itself (Bock & Levelt 1994: 964; 365). Hence, for example, the sentences below in (25)-(27) need not annotate TOP formally. On the other hand, in the Russian sentence in (28) the overlay of TOP to OBL (as well as of FOC to SUBJ) must be formalised. Conversely, because TOP is an *overlay* DF function, it will not be expressed if there is no nonDF to overlay it to. So, for example, the Italian sentence in (20) has neither TOP nor FOC because neither SUBJ nor OBJ is expressed syntactically – that is, a head-marking language such as Italian can mark them both on V by morphological means, which are part of the lexical entry (i.e., part of the word, rather than the sentence).

- (20) *lo bevo raramente*
['(I) it drink rarely' = 'I drink it rarely']

The reader would thus appreciate that in LFG the DFs TOP and FOC, which have a formal definition in the grammar and are subject to explicit constraints, are quite different from the 'topic' and 'focus' notions in other frameworks that may treat them ambiguously as discourse (rather than exclusively sentence) elements.

The LFG formalisation of DFs plays an important role in solving the functional uncertainty of the clause-initial phrase. In (17), for example, *Consuelo* is simultaneously TOP and SUBJ by default, so there is no need to mark TOP redundantly because SUBJ is already a DF (as well as a GF). On the other hand, in (18) the phrase *in writing* is not SUBJ but an ADJ formally marked as TOP. In this particular case, functional uncertainty is constrained in the interpretation of the sentence because a PP in the c-structure clearly points to an ADJ (or at least a noncore OBL). Uncertainty is more likely to arise, for example, with the topicalisation of a core argument such as OBJ, where TOP_{OBJ} may appear early in the sentence as a bare NP in c-structure. For an Italian example, we refer to (19) in chapter 3; in that sentence, because the first element in c-structure bears both the DF as TOP and the nonDF as OBJ formally in f-structure, LFG's Extended Coherence Condition is satisfied¹⁶ (Bresnan 2001: 69).

There is a further crucial clarification we wish to make in order to understand what DFs are in LFG, and hence in PT. On the one hand, TOP and FOC are universal GFs; on the other, their expression varies crosslinguistically.¹⁷ When a declarative sentence expresses TOP, more configurational languages, such as English, and head-marking languages, such as Italian and Spanish, tend to mark it, by default, by placing it in the most prominent position in c-structure, namely the first in the sentence, and may then place FOC after TOP. And this is how TOP is operationalised for these languages in this volume.¹⁸ However, besides position in c-structure, other types of languages may prefer to express TOP by morphological, lexical and, notably, prosodic means, or by a combination of means. For instance,

16 Completeness and coherence are general well-formedness conditions on f-structure (Bresnan 2001: 63–69). Completeness requires that every GF designated by a PRED be present in the f-structure of that PRED. The Extended Coherent Condition requires that all syntactic functions be integrated appropriately into the f-structure. This means that a TOP or FOC function must be integrated either by *identification* with an integrated function (in the case of SUBJ or ADJ which have their own PRED) or by *anaphoric linking* to an integrated function (in the case of argument functions other than SUBJ, such as OBJ and OBL).

17 In LFG's work, DFs have been investigated linguistically by many scholars (e.g., Bresnan & Mchombo 1987; Kroeger 1993), and formally implemented as the mapping of c-structure onto f-structure by Bresnan (2001), Dalrymple (2001), and Falk (2001). They are also sketched in Levelt (1989).

18 In Italian, for example, TOP may be encoded also in nondefault positions (i.e., non-initial) in c-structure (e.g., *li ha mangiati Pierino i fichi*, 'it was Pierino who ate the figs'). In this case, however, sentences are highly marked pragmatically and display specific prosodic contours. Since no learner considered in this volume produces any of them, they will not be further mentioned.

Japanese and Korean express TOP primarily by morphological means, namely the *-wa* marker and the *-un* marker respectively (for Japanese, cf. ch. 4). In interrogative sentences, which must express FOC, this DF can be expressed by prosodic, lexical (e.g., *wh*- words), syntactic and/or morphological means. More about FOC will be said when questions are treated in language-specific chapters because other important issues are at play here besides configurationality and head/dependent marking, such as whether questions are polar or constituent ones, and whether the latter are fronted or in situ in the specific language. As the reader will appreciate over the remainder of this chapter and the following ones, DFs play a key role in our understanding of the learner's syntactic development when they are used together with Levelt's processing principles, such as the attribution of prominence to an argument function, or to an ADJ – cf. our Prominence Hypothesis in § 4.2.1, (34)-(35).

In sum, (21) shows how GFs in LFG are subdivided into two major dichotomies: argument and nonargument functions on the one hand, and discourse and nondiscourse functions on the other.

(21) *Grammatical functions and their subdivisions (after Bresnan 2001: 96-97 and Falk 2001: 60)*

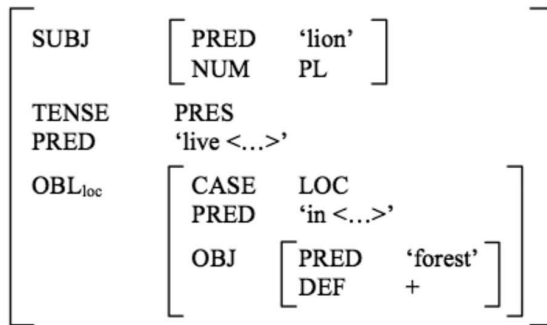
discourse fn		nondiscourse fns					discourse fns	
argument fns					nonargument fns			
core fns			noncore fns					
SUBJ	OBJ	OBJθ	OBLθ	COMP	XCOMP	ADJ	FOC	TOP

Let us now turn to the second type of grammatical information needed to interpret the sentence semantically, that is, the information conveyed by grammatical properties (or features). As we have seen earlier, these properties are part of lexical entries, and include diacritic features such as number, person, gender, definiteness, case, and tense, which all have their own values: in English, for example, singular and plural for number (e.g., *cherry*, *cherries*); first, second, etc. for person (e.g., *I*, *you*); masculine and feminine, etc. for gender (e.g., *gentleman/he*, *lady/she*); definite and indefinite for definiteness (e.g., *the banana*, *a banana*); nominative, accusative, etc. for case (e.g., *he*, *him*); and present, past, etc. for tense (e.g., *sing*, *sang*).

Thus in LFG, grammatical information in f-structure is represented by a set of attribute-value pairs; that is, given a particular f-structure, each attribute is always assigned a specific value. There are three types of values: (a) atomic symbols,

e.g., SG for singular; (b) semantic forms, e.g., love <x, y>, which stands for a kind of activity involving two arguments; and (c) f-structures, which themselves consist of attribute-value pairs. The f-structure for the sentence *lions live in the forest* is illustrated in (22).

(22) *F-structure for lions live in the forest* (Bresnan 2001: 46)¹⁹



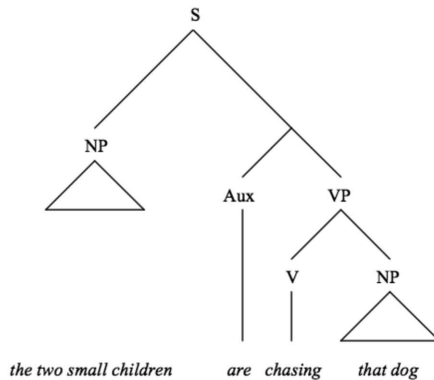
Turning now to c-structure, this is the overt expression of the functions and features that make up a syntactic expression (Falk 2001: ch. 2). It encodes three types of information: (a) word order, (b) constituent boundaries, and (c) the syntactic category of each word and constituent in the sentence – that is, whether a word is a noun, a verb, an adjective, etc., and whether a phrase is NP, VP, AP, etc. It is the level of representation of phrase-structure trees.

In contrast to f-structure, which encodes the invariant (universal) aspects of grammar, c-structure encodes properties that vary a great deal across languages. In this regard, we follow Bresnan (2001) in comparing, for example, English with Warlpiri, an Aboriginal language spoken in northern Australia. In English, c-structure is strictly organised, both linearly and hierarchically. A sentence can thus be made up of identifiable constituents, such as NP or VP, which are placed in specific positions; and GFs are encoded in c-structure configurations, with SUBJ outside of the VP and OBJ inside. Languages such as English are called configurational languages. On the other hand, in Warlpiri GFs are not encoded in c-structure; c-structure is flat, and all arguments are sisters of V. Thus word order is free, “so long as the auxiliary tense marker occurs in the second position” (Bresnan 2001: 6). Languages such as Warlpiri are called nonconfigurational languages. An English phrase structure (hierarchical and endocentric) and a corresponding Warlpiri one

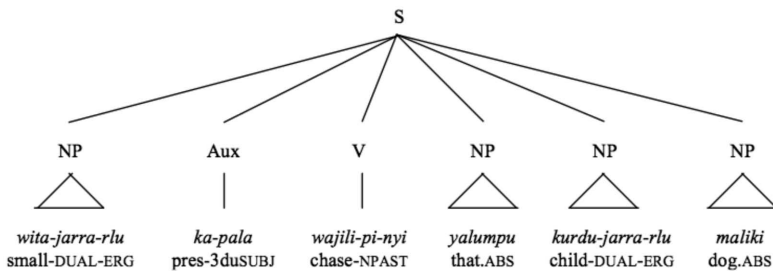
¹⁹ The internal structure of PREDS is omitted here, as in the original.

(flat and lexocentric) are illustrated in (23)-(24) respectively.²⁰ This alternative mode of c-structure organisation, *lexocentricity*, “associates syntactic functions directly with features borne by words rather than with the configurational relations of phrases in syntax” (Bresnan 2001: 109).

(23) *Phrase structure of the English sentence the two small children are chasing that dog* (Bresnan 2001: 5)



(24) *Phrase structure of the Warlpiri sentence glossed in English as the two small children are chasing that dog* (Bresnan 2001: 6)



20 The c-structure representations in (23)-(24), as well as others in (25)-(28) in this chapter, keep formalisation to a minimum in order to increase comprehension of the gist of our general presentation for readers who may be unfamiliar with the fine details of LFG annotations. Needless to say, we are aware that there is a great deal of variation in c-structure organisation across languages, and that in many of them what is here marked generally as a sentence (S) – as it is in Bresnan (2001) as a matter of fact – corresponds to IP. Dealing with specific languages, the following chapters will provide more precise and updated LFG formalisation.

In the Warlpiri sentence in (24), the actual word order is ‘two-small are chasing that two-children dog’, with the NP_{SUBJ} split by the V complex ‘are chasing’ and ‘that’, and with ‘that’ referring to ‘dog’ rather than to the adjacent ‘two children’. So, it is clear that (in Warlpiri, as in other nonconfigurational languages) the coherence of a conceptual unit is indicated by means of word *shapes* rather than word *groups*. Noncontiguous words that form a conceptual unit must share the same formal endings marking agreements. Indeed, the richness of the inflectional endings marking relationships between words and groups allows for great permutation of words and phrases in the sentence according to the speaker’s discourse-pragmatic needs. Such ‘sharing’ of inflectional features is the basis for the exchange of information that allows for feature unification. In terms of Levelt’s Model for language production, the greater the distance (in terms of syntactic levels²¹) between the words needing feature unification, the higher the cognitive cost of unifying them.

Typological variation between configurational and nonconfigurational languages sets up “competition between words and phrases expressing the same f-structure information” (Bresnan 2001: 101-102). This means that morphology and syntax play complementary roles, in the sense that morphology-rich languages show preference for lexical over syntactic expression for grammatical encoding, and vice versa: morphology-poor languages show preference for syntactic over lexical expression. However, as Bresnan (2001: 132) is quick to remark, along the typological continuum from strictly configurational to strictly nonconfigurational languages, natural languages may freely mix modes of organisation.

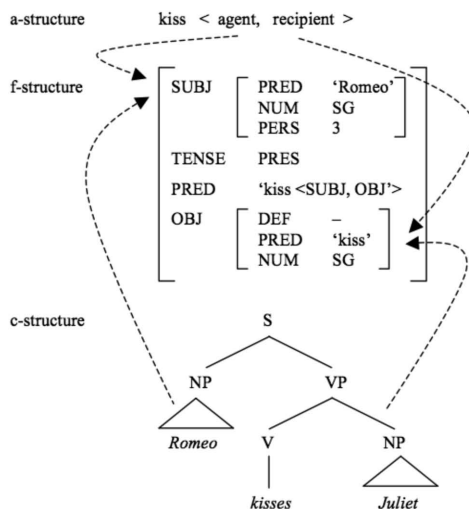
So far we have presented a-structure, f-structure and c-structure separately. This is possible, because each of these levels of representation is independent from the others, in the sense that they exist simultaneously and none is derived from any other. However, because each provides only a partial description of a sentence, it is important to specify the correspondence, or mapping, between the elements of these parallel structures. According to LFG, the main problem of a syntactic theory is to “characterise the mapping between semantic predicate-argument relationships and surface word- and phrase-configurations by which they are expressed” (Kaplan & Bresnan 1982: 174). Thus LFG grammatical formalism is essentially based on the correspondence architecture with which a sentence maps a- and c-structures onto the grammatical relations and properties in f-structure. Because each structure has its own primitives and hierarchy, the mapping between them can align in more than one way (Sells 2001), both across languages and

21 As Asudeh & Toivonen (2011) point out, even such extreme nonconfigurationality as found in Warlpiri appears to preserve some kind of ordering principle; e.g., in this language, AUX must be in second position.

within a specific language. For example, on the one hand, the default correspondences between the SUBJ function and the arguments can vary typologically along several dimensions; that is, in active languages like English, SUBJ maps by default onto the semantically most prominent role in a-structure; in accusative languages like Japanese, onto the argument in control of the eventuality; or in ergative languages like Dyirbal, onto the argument most affected by the eventuality. What all these possibilities have in common is the hierarchical prominence of the SUBJ argument on the selected dimension compared to other arguments (Bresnan 2001: 95). On the other hand, the SUBJ function takes no single universal form. Expression of SUBJ includes the NP in a certain phrase-structure configuration, as in configurational languages like English (cf. ch. 2); verbal inflection morphology, as in head-marking nonconfigurational languages like Italian (cf. ch. 3); and nominals bearing a specific case, as in dependent-marking nonconfigurational languages like Japanese or Russian and Serbian (cf. chh. 4, 5 and 6 respectively).

The mapping among the three structures is unmarked (or default) when the most prominent thematic role (i.e., agent) is encoded in the highest available GF (i.e., SUBJ) and occupies the most prominent position in c-structure, that is, the first position (Choi 2001). Levelt (1989: 266) calls this “congruent encoding”. An illustration of such mapping convergence is provided in (25) for the sentence

(25) *Canonical correspondences of a- and c-structures onto f-structure for the sentence Romeo kisses Juliet*



Romeo kisses Juliet, where the most prominent argument in a-structure, the agent, and the most prominent constituent in c-structure (the one in first position), both map onto the most prominent function in f-structure (the SUBJ). Less prominent thematic roles, if required, link onto less prominent functions in less prominent positions.

However, correspondences among the three structures can vary a great deal, and be more or less marked. In any language, for a variety of pragmatic reasons, the same propositional content can be expressed taking different perspectives, as we have seen in § 2.1. Such perspectives trigger different structural realisations. In most languages, sentences may vary between declarative and interrogative, between active and passive, and so on. Speakers may also choose to place a constituent in prominent position by topicalising or focusing it, or they may choose not to do so. Many of these structural choices are devices for directing the hearer's attention (Levelt 1989), and contribute to the representation of meaning, making communication more effective. However, how and how often these devices are deployed is language-specific.

Correspondence between the elements of the three levels of representation can be between arguments and GFs (mapping of a-structure onto f-structure) or between constituents and GFs (mapping of c-structure onto f-structure). The technical, formal details of these linking rules are complex, and well beyond the scope of this chapter (for their formalisation, cf. Bresnan 2001; Dalrymple 2001). We will illustrate in (26)-(27) one example for each type of correspondence, bearing in mind that GFs are here considered the 'relators' of c-structure to a-structure.

With regard to mapping a-structure onto f-structure, LFG proposes the Lexical Mapping Theory (Bresnan 2001: ch. 14; Dalrymple 2001: ch. 8, Falk 2001: ch. 4), which systematically explains how the conceptual representation of the thematic roles, represented by a-structure, is mapped onto GFs. In (25) we have seen an example of how this linking is predictable. But the eventuality described there can be realised differently, if speakers wish to establish a different perspective or point of view on the event they intend to communicate. For example, they may express the matter from the recipient's point of view, and prioritise *Juliet* over a demoted *Romeo*. This can be done by choosing a different lexical item, such as an intrinsically 'exceptional' V or an alternative 'nonbasic' V form. Let us clarify these two expressions, both borrowed from Pinker (1984).

Exceptional Vs are such because they "fail to respect canonical correspondences between thematic roles and grammatical functions" (Pinker (1984: 300). *Receive* is such a V. Other common exceptional Vs in English describe a psychological state or reaction, and include *please*, *delight*, *bore*, and *bother*. So, the eventuality of Romeo kissing Juliet, as well as canonically by the V *kiss* in (25), can be expressed noncanonically by means of the exceptional V *receive* (*a kiss*) in (26). On the other hand, it can be expressed also by a nonbasic V form, in this case the passive V *be*

given, as in (27). Like exceptional Vs, also nonbasic V forms fail to respect canonical correspondences between thematic roles and GFs. So what is the difference between them? Without venturing into the intricacies of lexical-semantics, it may be said that nonbasic V forms require nondefault mapping of a-structure to f-structure and may be morphologically derived from a more basic form which exhibits default mapping. For instance, in English the (active) V forms *give* and *scold* are morphologically related to the alternative (passive) forms *be given* and *be scolded* respectively, or in Italian the V forms *lavare* ('wash') and *abbracciare* ('embrace') are morphologically related to their reflexive and reciprocal nonbasic alternative forms *lavarsi* ('wash oneself') and *abbracciarsi* ('embrace each other') respectively. Exceptional verbs also require nondefault mapping, but their alternative forms, if available, are not morphologically related. For instance, in English *receive* may be a nondefault mapping V alternative to the default mapping V *give*, like in Italian *insegnare* ('teach') may be to *imparare* ('learn'). Another exceptional V in Italian is *pentirsi* ('repent'), but in this case, despite the *-si* form, no more basic form *pentire* exists.

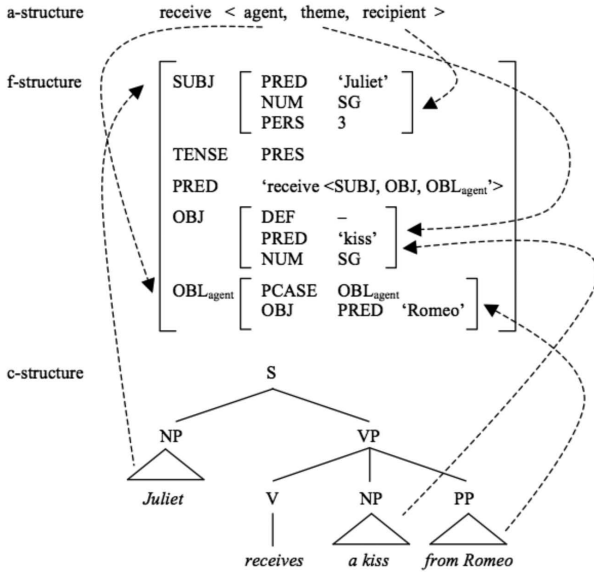
In any case, with both exceptional Vs and nonbasic V forms, it is important to note that – whether or not alternative forms are available, and whether or not alternatives are morphologically related – current LFG considers them all as separate lexical V entries which select their own set of arguments. In essence, a phenomenon such as passivisation “is a lexical relation change, not involving syntactic transformation, in that it can feed lexical processes of derivational morphology” (Bresnan (2001: 30). In fact, LFG follows a standard assumption in morphology: word formation processes such as derivation, compounding and conversion are morphological and not syntactic, and are thus the inputs to those processes. Consequently,

in many languages passivization, causativization, and other relation changing processes are inputs to lexical processes of derivational morphology such as nominalization, adjective formation, and compounding. It is not just the verb *forms* that are input to the lexical processes, but all their attendant syntactic effects such as changes in transitivity. (Bresnan 2001: 30, original emphasis – cf. also pp. 31-36)

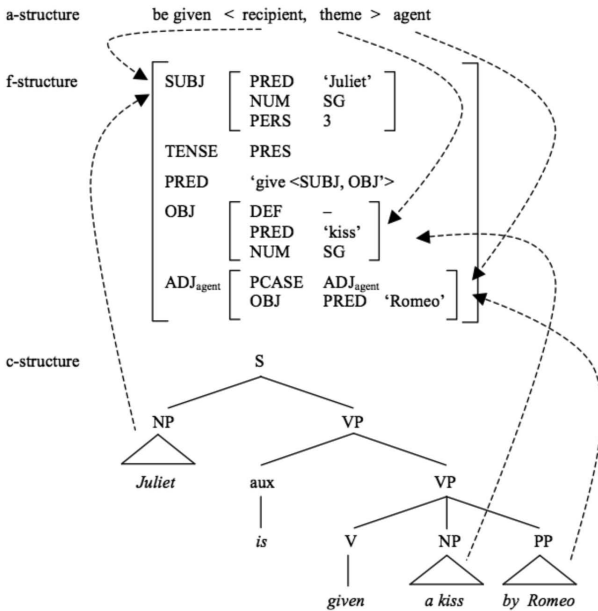
This is fully compatible with Bock & Levelt's (1994) and Levelt, Roelofs & Meyer's (1999) description of the nature and internal structure of the lexicon and its role in the formulation of the message in speech processing (cf. § 2.1 above).

Going back to our examples in (26) and (27), in either case, nondefaultness is due to the fact that it is the recipient, a less prominent thematic role than the agent, that is linked to SUBJ, the most prominent GF. The difference between *receive* in (26) and *be given* in (27) is that the passive V suppresses the agent, which then may be optionally mapped as ADJ (Bresnan 2001: 310). Yet notice that, in either case, with regard to c-structure, both sentences exhibit basic, canonical word order pattern – which, in the case of English, is SVO.

(26) *Nondefault mapping of a-structure onto f-structure for the sentence Juliet receives a kiss from Romeo*

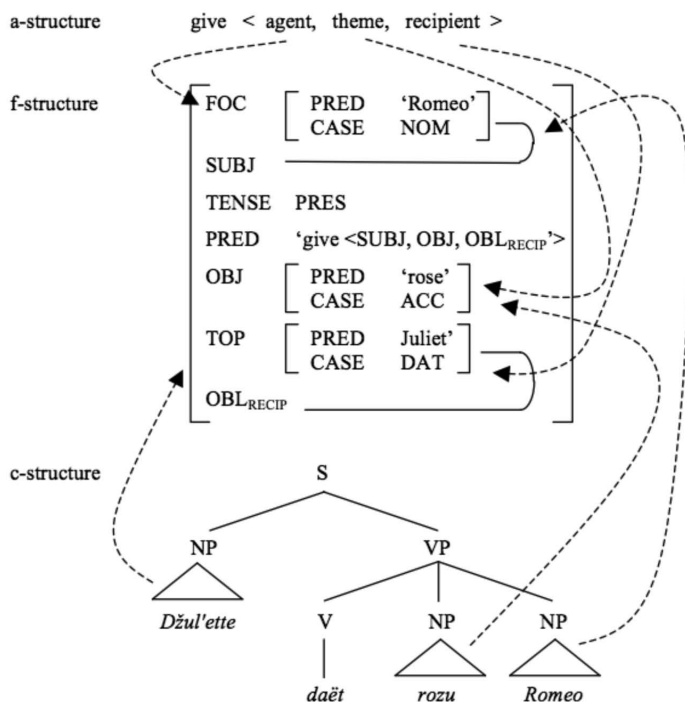


(27) *Nondefault mapping of a-structure onto f-structure for the sentence Juliet is given a kiss by Romeo*



With regards to mapping c-structure onto f-structure, we reiterate that, whereas f-structure functions are largely universal, c-structure configurations are language-specific. Language specificity is manifested in two ways. First, all languages have their typical (unmarked) canonical word order for core functions. For example, canonical order is SVO for English and Italian, SOV for Japanese, and VSO for Moroccan Arabic. Secondly, languages can be placed in different positions along the typological continuum from configurational to nonconfigurational languages, as we have just seen with the examples of English and Warlpiri in (23)–(24) at opposite ends. Among the European languages, Russian is less configurational than English, so we will use it to illustrate an example of noncanonical word order. In the eventuality of Romeo giving a rose to Juliet, if Russian speakers wish to assign prominence to *Juliet* (or better, *Džul'etta*) over *Romeo*, besides choosing the exceptional V *polučat'* ('receive') in a similar way to English, they can also choose to topicalise OBL and focalise SUBJ, as in (28). This involves a marked correspondence between c- and f-structures, with *Juliet* realised as TOPOBJ-RECIP preverbally, and

(28) LFG: noncanonical correspondences between c-structure and f-structure for the Russian sentence *Džul'ette daët rozu Romeo* [to Juliet gives a rose Romeo]



Romeo as SUBJ postverbally. Thus word order is noncanonical, and the argument function SUBJ is no longer associated by default with the DF TOP. On the other hand, notice that the mapping of agent as SUBJ and recipient as OBL remains canonical.

In concluding this brief presentation, we may summarise LFG as a lexically driven, psychologically plausible grammatical theory which provides an architecture for describing typologically diverse languages in a formal way. LFG provides PT with two fundamental concepts, ensuring that the different parts of a sentence actually do fit together:

- (29) a. the different syntactic levels – i.e., lexical level, phrasal level or sentence level – within or across which their elements require unification or merging of diacritic features and values, a process which iconically reflects performance; and
- b. the correspondences among a-, c- and f-structures; or more precisely, the ways in which elements in a-structure map onto those of f-structure, and elements in c-structure map onto those of f-structure.

The first of these two concepts, already incorporated into Levelt's 1989 Model, provided Pienemann in 1998 with the means to describe the learner's progress in the development of obligatory morphological and morphosyntactic structures (cf. § 4.1). Then in 2005, thanks to the second of these LFG's key concepts, Pienemann, Di Biase & Kawaguchi began to focus more sharply on optional syntactic constructions afforded by the alternative ways in which languages handle mapping their c-structure and their a-structure onto f-structure (cf. PT's 2005 Topic Hypothesis, which we propose to expand into the Prominence Hypothesis in § 4.2.1, and the Lexical Mapping Hypothesis in § 4.2.2). We should note, in concluding this part, that DFs and Lexical Mapping Theory were 'officially' formalised in LFG around 2001, that is, after two decades or so of research into many typologically disparate languages, including Australian Aboriginal languages. Yet, Levelt (1989) had already incorporated into his processing Model of the speaker considerations (a) about the default way in which, for example, the topical fragment available early from the conceptualiser corresponds to the earliest retrieved lemma(s), which in addition may receive primacy in grammatical encoding and is commonly generated as SUBJ NP in sentence-initial position, and (b) about how the speaker's perspective may direct attention and establish prominence and/or speaker's perspective otherwise. Furthermore, we should also note that the hierarchies of LFG's parallel structures (a-, c- and f-structures) iconically represent cognitive processes, and as such lend themselves to supporting processing interpretations of speaking.

3. PT's key concepts

On the basis of its feeder theories, PT describes, explains and predicts the development of morphology and syntax for any typologically different L2 by focusing on the development of the processing procedures (described by Levelt's Model) required for the production of L2 structures (described by LFG).

One of PT's central claims in Pienemann (1998) is that the sequence with which learners develop their grammar is not at all arbitrary but follows the time course with which the grammatical encoding of the lexicon unfolds in Levelt's Model. Hence the language processing sequence described for L1 mature speakers in (11) above foreshadows the developmental progress described for L2 learners in § 4 below. The sequenced activation of the processing procedures allows for the production of language structures, first those structures that do not require any exchange of information among constituents, later on those that do require it at the phrasal level, and finally at the sentence and higher levels. Exchange of information is a key concept here.

PT, then, spells out the hypotheses for the developmental sequences of L2 morphology and syntax in learners' interlanguage. That is, if learners are able to apply processing procedure x , they will be able to produce morphological or syntactic structure y using procedure x . Subsequently, if learners are able to use processing procedure x^n to produce structure y^n using that x^n procedure, it means that they are also able to use the preceding procedure x and produce structure y accordingly; that is, the process of building up procedures and corresponding structures is cumulative. Implicational hierarchy is a key concept here.

Based on the activation of implicationally arranged processing procedures, PT conceives L2 acquisition in terms of sequential progression through a series of stages. For morphology, progression is operationalised in terms of feature unification, and measured by the syntactic level on which lie the elements requiring feature unification, that is, exchange of grammatical information, in the target language (Pienemann 1998). For syntax, it is operationalised in terms of word order in *c*-structure, and the degree of flexibility in the mapping of *a*-structure onto *f*-structure measured by the canonical/unmarked vs noncanonical/marked sets of correspondences among these structures (Pienemann, Di Biase & Kawaguchi 2005).

Implicit in the amount and type of information that needs to be grammatically encoded, and then finally produced, is the cognitive cost of temporarily storing in the syntactic buffer the bits of partially encoded information that will be later required for further encoding. Processing cost is a key concept here. The more information needs to be exchanged, and the longer it needs to be kept available in the syntactic buffer of working memory, the greater the processing

cost for the learner.²² Crucially, the more costly the encoding of the structures, the more difficult the learning, and the later these structures will develop in the learners' interlanguage.

However, the cost of grammatical encoding to learners decreases as processing procedures become automatised through frequent activation. Thus the learners' progress depends on both the ability to activate new procedures along their implicational sequence, and the automatising of previously emerged components. In the meantime, while more advanced procedures are not yet available and earlier ones not yet automatised, the least costly solution for learners is to resort to default (or unmarked) structures involving the simplest one-to-one relationship between form and function, as long as there are words in the lexicon that match the conceptually generated message. Defaultness – (un)markedness, or canonicity – is yet another key concept in understanding PT.

4. The learner's progress

Having briefly summarised PT's theoretical bases in §§ 1-2 and key concepts in § 3, we now look at how these shape the learner's progress along the development path.

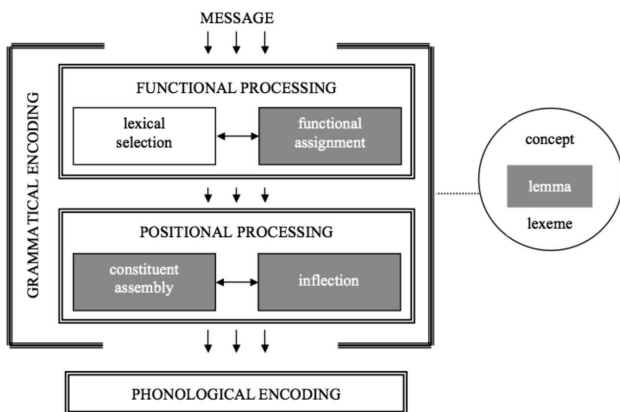
Briefly stated, PT hypothesises that, in converging towards the target language, L2 learners initially encode words and formulas phonologically but not yet grammatically. That is, they are produced in the order in which they become available in the conceptualiser, as long as the lexicon stores lemmas (in whichever state) that match (or approximate) the preverbal message. Grammatical encoding begins with words exhibiting minimally categorial features, and conceptual structure mapping onto surface structure canonically according to "prominence hierarchies" (Choi 2001: 24). This corresponds to Levelt's notion of "congruent grammatical encoding" (1989: 266), which entails the mapping of the highest role in the thematic hierarchy (the agent) onto the highest GF (the SUBJ) in the most prominent c-structural position (the first in the clause). Grammatical encoding develops further when words acquire more features, and learners become able to process discourse-pragmatic requirements that require noncanonical correspondences among the three hierarchies. A case in point is a passive structure, where a thematic role lower in the hierarchy (e.g., the patient) maps onto SUBJ); another is topi-

²² For the native speaker hardly any cost is involved here because this process is proposed by Levelt (1989) to be largely automatic and hence requires no attentional resources (cf. Poullisse 1997: 204).

calisation, when a GF lower than SUBJ (e.g., OBJ) is placed in the most prominent c-structural position (first in the clause).

In (30) we reproduce the components of grammatical processing already illustrated in (4) as fully activated in mature native speakers. Here, a grey shadow in the relevant components shows graphically the early learners' limited annotations in the lexicon and their limited ability to assign functions to thematic roles and constituents, and values to features. Learning progresses as the L2 formulator is built up (cf. De Bot 1992), and these areas become gradually clearer and more actively involved. The task of building up an L2 formulator – or indeed the very existence of a totally different formulator for each language – may vary according to the learners' L1, as De Bot (1992) notes. If the two languages are typologically unrelated the task is more arduous and the learning slower. If they are closely related, it may even be doubtful whether a whole new formulator is needed, and in any case many of the categories and procedures needed for speech may already be in place and operative. This would result in faster learning.

(30) *Levelt's Model: components of grammatical processing at an early stage of L2 acquisition (after Bock & Levelt 1994: 946)*



Although the sequences in the parallel learning of morphology and syntax certainly interface in important ways (cf. § 4.3), we present them separately in §§ 4.1 and 4.2 respectively.²³ In the introduction to this part I of the volume we have already

23 There is some evidence across languages for L2 syntax to emerge, consistently, before morphology. Bonilla (2014) for instance found that syntax emerges before morphology in all her Spanish L2 learners.

stated that the reason for doing so is that the two developmental sequences, as proposed for PT in Di Biase & Kawaguchi (2012), depend on two different sets of motivations. On the one hand, we have the original psycholinguistic procedures of Kempen & Hoenkamp (1987) and Levelt (1989) modeled in LFG by the mechanisms of feature unification, and adopted for PT in Pienemann (1998). On the other hand, we have the different kinds of correspondences among the three parallel LFG a-, c-, and f-structures, adopted for PT in Pienemann, Di Biase & Kawaguchi (2005).

In our presentation of both the universal schedules in this chapter and the language-specific ones in the following chapters, we prefer to avoid the use of numbers for identifying stages, unlike most tables presenting PT sequences in the considerable volume of PT literature since Pienemann (1998). Two reasons guide us in this decision. First, although conveniently synthetic, numbers are not used consistently, especially whenever authors feel the need to highlight stages within a stage. This is particularly evident when, in languages such as English, the activation of phrasal procedure includes the emergence of both NP morphology and VP morphology, whose relative sequence is not clear (cf. ch. 2, § 2). This may not be relevant in languages without VP. Secondly, as well as the sequence for morphological development, PT has now added two other sequences for syntactic development, as we shall see in § 4.2 in this chapter. Because the correspondences among these three parallel lines of development are very much an empirical issue crosslinguistically, and may well turn out to exhibit differences, these should not be pre-empted. Thus the conveniently synthetic use of numbers for stages may become cumbersome if it were necessary to specify which of the three sequences was actually being referred to, and whether or not it may correspond to the same-numbered stage across lines of development.

4.1. Morphological development

PT hypothesises that the availability of increasingly more demanding processing procedures defines the learners' progress through a sequence of stages which depend on the increasingly greater syntactic distance (i.e., in terms of hierarchical levels) between the linguistic elements requiring exchange of information for their appropriate grammatical production.²⁴ This postulated universal sequence is shown in (31). Of course what exactly is unified over what syntactic distance must be defined language-specifically. Once the linguistic exponents are hypothesised – that is, what language-specific structure requires what procedure (belongs to what stage) – then these hypotheses are tested on spontaneous speech produced by learners of the

²⁴ Franck, Vigliocco & Nicol (2002) managed to separate experimentally the role of syntactic distance vs linear distance in sentence processing.

particular language. Language-specific sequences and examples will be presented and discussed in §§ 2.1, chapters 2, 3 and 4 for English, Italian and Japanese respectively. Sequences for other languages (German, Russian, Serbian and Spanish) are also hypothesised, though less extensively, in subsequent chapters.

(31) *PT: hierarchy of processing procedures – morphological development (after Pienemann 2005b: 14)*

STAGE	T1	T2	T3	T4	T5
S-BAR PROCEDURE	–	–	–	–	interclausal information exchange
SENTENCE PROCEDURE	–	–	–	interphrasal information exchange	+
PHRASAL PROCEDURE	–	–	phrasal information exchange	+	+
CATEGORY PROCEDURE	–	lexical form variation	+	+	+
LEMMA ACCESS	words & formulas	+	+	+	+

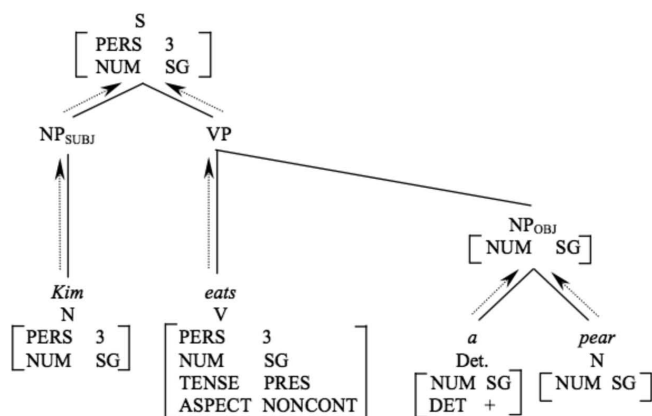
Initially, as soon as one or more words are appropriated by the learners, the only L2 procedure they can activate is lemma access, which is sufficient to allow the production of the word. Unable to activate any further grammatical procedure, with regard to morphology they produce only formulaic expressions or single words with no overtly meaningful formal variation. The main reason for this inability is that, at this early stage, the L2 word or formula is learned as a whole and is not further analyzed. Or, putting it in another way, and paraphrasing Bock & Levelt's (1994) and Levelt, Roelofs & Meyer's (1999), the three-level system of the mature L1 speaker's lexicon, represented in (2)-(3) above, is reduced to two out of the three levels: the conceptual level (meaning) and the lexeme level (sound).

At the next stage, learners begin to annotate their lexicon, and develop a system of lemmas whereby lexical concepts acquire a syntactic category and later its characteristic diacritic features. This feeds the process of syntacticisation and activates the category procedure to grammatically distinguish, for instance, Vs from nonVs. The values distinguished earlier by learners are usually those of diacritic features expressing the more transparent conceptual representation with phonologically consistent and possibly frequently occurring forms – such as plural marking for Ns, or aspect/tense for Vs (in languages that do mark such differences). At this

category procedure stage, then, overtly meaningful formal variation begins to emerge but the grammatical information thus annotated does not carry beyond the word level. In the V lemma, in languages such as English for example, tense and aspect markers remain contained within the word and may serve more as categorial markers to differentiate Vs from Ns than as actual markers of tense, aspect (cf. Johnston (1994) for a discussion of the categorial role of *-ing* in early ESL learners). This means that the relevant diacritic feature is available in the same location where the morpheme for the marking of aspect (*-ing* in English, for example) must occur. Because there is no exchange of information taking place, nothing is stored for further use somewhere else in the sentence. For those lexical entries which are not yet fully annotated, the learner will likely use default or citation forms, the least marked and most available in the input, such as (but not necessarily for all lexemes) the singular form in languages that mark number, the nominative form in those that mark case, non-past forms for tense, and so on.

With the next step forward learners reach the phrasal procedure stage. If the category of the head lemma is N, then the NP procedure can be called in order to produce an NP. If a determiner or modifier is added, the value of the head plays a key role. That is, the grammatical information of the head lemma N must be temporarily stored in the NP procedure to be checked against that of the other constituent(s) within the phrase. In order to do this, information must be exchanged among the words that in the target language require feature unification. For example, in the English phrase *a pear* in (12) above – here reproduced in (32) with the addition of arrows showing the path/process of the unification –, DET and N share the feature number, and require feature unification (represented at the NP node in the LFG diagram), which in this case concerns the value singular.

(32) *An illustration of processing hierarchy for Kim eats a pear showing phrasal and inter-phrasal procedures*



At the sentence procedure stage, the phrase needs to be attached to the S(entence)-node (i.e., the mother node in the tree structure – although, notice that S is not necessarily the appropriate mother node in all languages or for all sentence structures; in English, e.g., it is typically IP – cf., e.g., Dalrymple 2001: 53–54, 60–61; Falk 2001: 40), with the sentence procedure ensuring, in any case, the functional destination of the NPs associated with the argument roles of the V, such as NP_{SUBJ} or NP_{OBJ}. Here again, the required information relating to a phrase's values must be stored until the diacritic feature is assigned to the appropriate place elsewhere the sentence and the values checked for compatibility. For example, in the English sentence *Kim eats a pear* in (32), the NP_{SUBJ}'s number (SG) and person (3rd) values are kept in the syntactic buffer (short-term memory store) until the appropriate V form, with the bound morpheme *-s* is retrieved for the V *eat*. The S-procedure then checks the compatibility of the information coming from N and the V phrases. In this example, number and person come from the N *Kim* inside the NP_{SUBJ} and the V *eat*, with its person and number values, from inside the VP. As the arrows indicate, this information travels over different phrasal nodes first, is carried up to the S-node and then checked for compatibility, that is, the person and number values in the V must be compatible with those of the NP_{SUBJ}, for the required interphrasal exchange of information.

Further along the morphological developmental path, at the last stage, learners activate the subordinate clause procedure (if the target language requires one at all). They thus become able to exchange information about the values of relevant diacritic features between elements in different clauses related to each other by subordination (i.e., the main clause and the dependent one). For example, interclausal information exchange is required in the subjunctive mood reading of the rather rare and formal English sentence in (33), where the information exchanged is that the subordinate clause does not require feature unification between SUBJ and V (*Kim eat*), whereas the main clause does (*Ann suggests*).

(33) The doctor suggests that Kim eat less

In this volume the complex area of subordination will not be addressed. It is an important research gap in PT, which requires further theoretical elaboration and focused empirical investigation.

As we can see from the examples above, the exchange of information within and across constituent boundaries can entail unification involving grammatical operations of different nature. Unification includes morphological concord (e.g., *fiori-PL.MASC profumati-PL.MASC*, 'scented flowers', in Italian), semantic compatibility (e.g., *these books*; *three cats*; *many people* in English), and grammatical government (e.g., *has gone* vs *is going* in English; *s babuškoj-INSTR.*, 'with grandmother' in Russian).

4.2. *Syntactic development*

For syntax, as for morphology (cf. § 4.1), PT hypothesises a staged development (Pienemann, Di Biase & Kawaguchi 2005). This comes about as learners gradually develop from default, fixed solutions in linking arguments and constituents to GFs towards noncanonical, more flexible solutions. That is, they gradually learn to (a) annotate ever more detailed and specific lexical requirements and (b) grammaticalise discourse and pragmatic information. After the initial stage of single words and formulas, there are two parallel paths ahead for syntactic development, both of them leading to the enrichment of f-structure: on the one hand, after learners begin to map c-structure to f-structure in a canonical way, they start ordering constituents more freely, as required by discourse and pragmatic reasons. On the other, after they begin to map a-structure to f-structure in a default way, they then start mapping semantic roles more freely on GFs as required by exceptional lexicon and/or by discourse and pragmatic reasons.

In other words, one path develops the ability to assign GFs to constituents by anchoring them at first to a fixed position in the linear order of the clause (SUBJ in initial position then OBJ), and later on disengaging them from it – thus allowing, for example, the learner to place a GF other than SUBJ in first position. This path was originally formulated as the Topic Hypothesis in Pienemann, Di Biase & Kawaguchi (2005). However, two important problems arise from that formulation. One is the operationalisation of TOP as a GF different from SUBJ, and the other is the neglect of FOC, the third DF. After Mycock's (2007) work on constituent questions in the LFG framework and its deployment in recent work within the PT framework (e.g., Bettoni & Di Biase 2011; Di Biase & Bettoni 2015), it may be useful to recast the original Topic Hypothesis more generally, such that it embraces all three grammaticalised DFs. This would then include, besides TOP for the development of declaratives, also FOC for the development of interrogatives. In an attempt to account also for processing as a basic constraint within which speakers – and naturally enough, learners – construct their utterances from intention to articulation (to paraphrase Levelt), we propose to replace the 2005 Topic Hypothesis with our Prominence Hypothesis. In previous drafts of this volume, as well as in Bettoni & Di Biase (2011) and Di Biase & Bettoni (2015), this latter hypothesis was called the Discourse Functions Hypothesis. Its current fine-tuning as Prominence Hypothesis owes a lot to the stimulus provided by discussions with Gabriele Pallotti. In changing its label we realise, of course, that the Lexical Mapping Hypothesis may also be motivated by discourse concerns with prominence. However, as we will see, in the Lexical Mapping Hypothesis there is more at stake than just prominence. The Prominence Hypothesis deals, then, only with such prominence as achieved by precedence relations on c-structure alignment, that is, by adding the required discourse information in f-structure and the corresponding c-structure elements.

The second path is heralded by the 2005 Lexical Mapping Hypothesis, which traces the ability to assign GFs to thematic roles first in a default way (e.g., agent to SUBJ, theme/patient to OBJ) and later in a nondefault way. What motivates nondefault lexical mapping? Unlike the Prominence Hypothesis path, this path is primarily lexically motivated. So, non-default mapping is lexically required either by specific verbal predicates, such as intrinsically exceptional Vs or alternative non-basic V forms, to use Pinker's (1984) terminology, but recalling that it is not just the 'form' of the V entry to play a role but, primarily, its attendant conceptual-syntactic information (Bresnan 2001: 30), as we have seen in § 2.2 above. Then, of course, languages may well allow speakers to combine both types of resources; this would require operating both on word order in c-structure and on argument mapping, which entails a high level of development and automaticity. We will not go into this presently.

Some readers may have concluded by now that this treatment of syntactic development along the two schedules based on the Prominence Hypothesis and the Lexical Mapping Hypothesis makes one of the hypotheses in Pienemann, Di Biase & Kawaguchi's (2005) extension, namely, the Unmarked Alignment Hypothesis, redundant. This is, indeed, a desirable result in theory construction because it is more parsimonious to put forward fewer hypotheses to cover the same (or a greater) area. Furthermore, as we said in the introduction to this volume, the Unmarked Alignment Hypothesis actually conflates c- to f- and a- to f-structure mapping at the first grammaticalised stage following the single-word and formulaic one. This, however, may be misleading conceptually because – as Conroy (2007) remarks – it assumes that canonical word order in c-structure necessarily entails default mapping of thematic roles onto GFs, thus conferring an unintended universality to that particular convergence (i.e., agent-SUBJ mapping in first position, followed by the theme/patient-OBJ mapping). On the other hand, the Prominence Hypothesis and the Lexical Mapping Hypothesis account for their own respective paths independently from each other from the very beginning of the grammaticalisation process.

In the next two sections, we spell out the two hypotheses in further detail and illustrate the two developmental sequences respectively in (34)-(35) and (42)-(43). Then in § 4.3 we discuss the interfaces between them and with morphological development.

4.2.1. The Prominence Hypothesis

This hypothesis accounts for the staged grammaticalisation of the DFs TOP and FOC at the sentence level. When the activation of the category procedure takes place and distinguishes at least between Vs and Ns, Vs begin to acquire their pivotal role in the sentence frame, and participants in the eventuality begin to be

aligned according to the specific canonical word order of the target language, as learners follow positive evidence in the input. But, what is canonical order? Is there a universal one?

The notion of canonical order has been hotly debated at least since Greenberg (1963) – see Song (2012) for a historical and theoretical account. In a 402-language sample based on a database of 1063 languages, Tomlin's (1986) typological work found two prevalent linear orders for SUBJ, V and OBJ: SVO and SOV, with no statistically significant difference between the two orders (41.79% vs 44.78% respectively). The third order was VSO (9.2%) leaving the remaining three possible orders (VOS, OVS, OSV) with less than 5% between them. On the other hand, working on language genetic groupings (genera) rather than languages as single entities, Dryer (1989) found a preference for SOV, confirmed with an extended database of 1377 languages in the *World Atlas of Linguistic Structures*: 41% SOV languages, 35% SVO, 7% VSO and only about 3% distributed over the other three possible order, with the remaining 14% or so languages displaying no dominant word order (Dryer 2013). All these findings nevertheless confirm that in the great majority of languages SUBJ precedes OBJ. So the typology of word order would tend to support Keenan's (1979) *Subjects Front Principle*. And this principle can be further supported if we consider the order of precedence in the widely accepted GF hierarchy proposed by Keenan & Comrie (1977): SUBJ > OBJ > OBJ θ > OBL > ADJ (cf. also Choi 1999, Bresnan 2001, and § 2.2 above). In sum, excluding the position of V, on typological accounts (whether numerical or genealogical) and hierarchical accounts, SUBJ should be structurally more prominent than non-SUBJ GFs, and, in the absence of other arguments, OBJ will follow as the closest sister of V.

Canonical word order is then the optimal (and prevailing) order with which each language organises its basic constituents in the c-structure of the prevalent type of strings (i.e., simple, active, affirmative, declarative, minimally presuppositional, and pragmatically neutral sentences). Thanks to its very predictability, L2 learners can produce canonical word order with minimal processing cost, and minimally specified functional assignment.²⁵ That is, at an early stage functional assignment may be purely positional, with a nonhierarchical, flat structure. Not surprisingly, the L1 or L2 learner's initial reliance on the canonical word order of the target language is well attested by a large number of corpus-based longitudinal and cross-sectional studies of typologically different languages, outside of the PT framework (e.g., Bever 1970; Sano 1977; Slobin & Bever 1982; Clahsen, Meisel & Pienemann 1983; Pinker 1984; Johnston 1985a; Cook 1993; Sasaki 1998), and

²⁵ Readers may also note in this connection that all learners will already have the notion of SUBJ and OBJ in their L1 though with their own language-specific characteristics.

within PT (e.g., Di Biase & Kawaguchi 2002; Itani-Adams 2009; Kawaguchi 2005, 2011; Mansouri 2005; Medojević 2014; Pienemann & Håkansson 1999; Yamaguchi 2010; Wang 2010; Wirbatz 2014; Zhang 2005) – where L1-L2 transfer occurs it is developmentally moderated, as Pienemann, Di Biase, Kawaguchi & Håkansson (2005) show across a range of typologically different languages. We also assume that canonical order is language specific: for example, SVO in English and Italian, and SOV in Japanese.

We also admit that all languages studied so far exhibit the canonical SO order – and that it would be extremely interesting to test PT's hypotheses on OS languages, which are in any case very rare. Rather than the nature of canonical order itself, however, what we wish to stress here is that canonical correspondences between linear positions and GFs and DFs are typically learned before noncanonical ones.

The assumption that at the initial syntactic stage learners can produce canonical word order – whose default solution is to make the first NP the SUBJ/TOP, and the second NP the OBJ/FOC – would imply that the S-procedure is already operative in their interlanguage. This creates some confusion (cf. also Kempen's 1998 critical remarks), which Pienemann (1998: 87) tries to disentangle by labelling this canonical ordering principle a 'simplified' S-procedure. Without entering into detail about procedures, which crucially determine morphological development (cf. § 4.1), we make the point here that the kind of GFs emerging at this stage are primarily positionally determined. That is, although learners are now able to assign GFs to sentence constituents, they are unable to operationalise the full range of functional assignment exponents apart from position. In other words, at this first syntactic stage of development, they can produce only one frame: a canonical string mapping c-structure to f-structure in a fixed correspondence. Unable to mark GF/DFs by other means than position, the only frame learners can produce is fixed: SUBJ/TOP precedes OBJ/FOC.

In configurational languages such as English, position may be more or less the full story, but not so in nonconfigurational languages. The learner's gradual progress from functional assignment dependent on position in a fixed frame to functional assignment independent from position depends on the availability of the necessary resources (e.g., inflectional morphology). In this regard, it is useful to keep in mind the temporal sequence in which Bock & Levelt (1994) clearly indicate that functional assignment precedes inflection (cf. (29) above), and our interpretation of the interfacing between syntactic development and morphological development in § 4.3, as well as language specific exemplifications in part II of the volume.

Let us then move on to consider richer frames and more flexible word orders. In order to enhance their expressiveness, speakers may wish to give prominence to a particular entity in the eventuality they intend to communicate to their interlocutor. There are various ways in which they can do this, such as by prosodic, lexical or syntactic means (Levelt 1989: 266-67) in various combina-

tions. We would add here, morphological means (e.g., topic marking in Japanese, Korean, etc.). The way learners progress through the staged development of syntax – away from the rigidity of canonical word order towards the flexibility of optional choices allowed by their L2 – was spelled out by Pienemann, Di Biase & Kawaguchi in their 2005 Topic Hypothesis, which we now propose to replace with a more inclusive Prominence Hypothesis in (34). This is schematically summarised in (35).

(34) *Prominence Hypothesis*. In second language acquisition learners will initially not differentiate between grammatical functions (GFs) and discourse functions (DFs), for example, between SUBJ and TOP. Differentiation begins when an element such as an XP, or other lexical material, is added to the canonical string in a position of prominence in c-structure, that is, the first in the sentence. This element may be TOP in declaratives or FOC in interrogatives leaving, crucially, the canonical string unaltered. At the next stage, learners will be able to construct noncanonical strings assigning prominence to any constituent in an unequivocal way.²⁶

(35) *PT: syntactic development based on the Prominence Hypothesis*

STAGE	STRUCTURES
NONCANONICAL WORD ORDER	TOP _{XP} marked orders
	FOC _{XP} marked orders
XP _{DF} CANONICAL WORD ORDER	TOP _{XP} SVO / SOV / ...
	FOC _{WH} SVO / SOV / ...
CANONICAL WORD ORDER	SVO / SOV / ...
	[QUE ^P SVO / SOV / ...]
LEMMA ACCESS	single words; formulas
	[QUE ^P single words; formulas]
QUE ^P = the QUE feature is realized by prosody alone	

Our formulation of this hypothesis represents a significant expansion on its 2005 version, in so far as (a) it incorporates a processing principle explicitly, that is, the

26 This formulation of the Hypothesis may not account for all aspects of languages with in-situ constituent questions, an issue which should be resolved empirically.

fact that the sentence-initial position has higher prominence; and (b) it coherently includes interrogative as well as declarative sentences, and may thus be applied to a wider typological range of languages.

But what exactly may the learners be learning when attributing prominence? Here we follow Choi (1999: 133), who uses i-structuring constraints operating on c-structure to identify the contextual (discourse) dimensions of novelty and prominence, and marks each of them by a binary feature, respectively [\pm NEW] and [\pm PROM], as shown in (36).

(36) *Information structuring constraints (i-structure/c-structure correspondence)*

NEW: [–New] should precede [+New]
 PROM: [+Prom] should precede [–Prom]

These feature-value pairs would enable the identification of SUBJ in its default initial position as [–New] and [+Prom] preceding a default OBJ [+New] and [–Prom], whereas a topicalised OBJ would be assigned [+Prom]. So the informational status of the V arguments in a given f-structure, characterised by their informational feature structure, should predict the optimal alignment of the arguments in c-structure. Readers may recall that the f-structure does specify functional assignment and, if necessary, the overlaid DFs (TOP, FOC), but f-structure is unordered, hence underspecified in terms of alignment. If the arguments do not differ in informational status, the canonical word order prevails, but if they do differ, the optimal structure should result from the appropriate ranking of the constraints. See Choi 1999, 2001 for a treatment of information structuring within Optimality Theory and LFG. The information status of the arguments cannot be derived on the basis of their referential (semantic) features alone, but also by the speaker's perspective. As Dalrymple & Nikolaeva (2011: 14, original emphases) put it, topicality, for instance, "is not an inherent property of the referent, and although it correlates with the role played by the referent in the preceding discourse [...] it depends on the speaker's assessment of its saliency within a given **communicative context**". They then assert that "[s]ince we view information structure as part of grammar, we treat grammaticality as including pragmatic well-formedness." An example of the different information status of the elements in the same eventuality is shown in (37). In this example by Choi (1999: 46-47), in answer to the *where* question in (37a), the FOC *on the table* is the only element marked with the [+NEW] feature in (37b-d). But why are there three different answers? The differences are due to the speaker's manipulation of prominence through the values of the [\pm PROM] feature.

- (37) a. where did you put the knife?
 b. I put [the knife] [on the TABLE]
 c. [the knife] I put [on the TABLE]
 d. [on the TABLE] I put [the knife]

	[I]	[put]	[the knife]	[on the TABLE]
	[−New]	[−New]	[−New]	[+New]
a.	[−Prom]	[−Prom]	[−Prom]	[−Prom]
b.	[+Prom]	[−Prom]	[−Prom]	[−Prom]
c.	[−Prom]	[−Prom]	[+Prom]	[−Prom]
d.	[−Prom]	[−Prom]	[−Prom]	[+Prom]

Let us now proceed to illustrate the Prominence Hypothesis summarised in (35) in a general way, mentioning only some representative structures of English. For more detail and language-specific schedules and examples, the reader is referred to the chapters in parts II and III.

At the very first grammatical stage after the single-word and formula stage, we include the possibility for learners to also produce questions. Constituent questions may be produced by placing the *wh*-word in situ, that is, in the same position as their declarative equivalents, irrespectively of whether their target language marks FOC in situ or clause initially. Polar questions are grammatically marked by prosodic means, represented in the table by a prefixed QUE (for the development of questions, cf. Yip & Mathews (2006), and chh. 2 and 8, this volume, for English and Italian examples respectively). Typical additions to declarative sentences at this stage are time and place specifications (cf. the notion of Stage Topic in Erteschik-Shir 2007), normally placed to the right of the clause, thus leaving unaltered the default association of the DF TOP with SUBJ, as in (38).

- (38) I eat lunch in park

The next stage of XP canonical word order comes about when, for discourse or pragmatic reasons, learners become able to add to the canonical string an XP (that is, a phrase of any category) in the sensitive initial position. Using English to exemplify, the DF assigned to this phrase will be TOP in declaratives or FOC in interrogatives, as in (39a-b).

- (39) a. in Libya the people not drink beer
 b. where you buy this?

This fronted addition plays a crucial role in the learner's development because it promotes a differentiation between the DFs TOP or FOC and the canonical string. That is, in declarative sentences the TOP function will now be assigned to this new

constituent, rather than to SUBJ by default. The less costly choice in declaratives is that this new constituent be ADJ rather than an argument of V. And indeed Pienemann, Di Biase & Kawaguchi (2005: 232) report that empirical studies of the development of a range of languages identify ADJ among the first nonSUBJ constituents to occur in sentence-initial position – but see chapter 9, this volume, for a study including both declaratives and questions. What matters most, at this XP canonical word order stage, is that, whereas earlier on the identification of SUBJ/TOP by (first) position on c-structure did not allow for the differentiation of SUBJ and TOP (and indeed there was no need to distinguish between them: if there is only canonical order, TOP is SUBJ), now the appearance of an XP in the most prominent position before the canonical string triggers a dislocation of SUBJ from its canonical first position which interferes with the close connection between SUBJ and TOP: the TOP function is assigned to the prominent constituent, which is followed by SUBJ and the remaining constituents in canonical order.

Although the fronting of a circumstantial element bearing the nonargument function ADJ (usually expressed with an adverb or a PP) may be common, learners may occasionally introduce in first position yet a different new element expressed with a bare NP, co-referential with one of the argument functions such as SUBJ and OBJ, as shown in (40a-b). Such topics, however, at this stage, are ‘external’ (cf. Bresnan 2001: 69) because the remaining part of the sentence (*they read the newspaper*) not only displays fixed canonical word order, but it is also complete and fully coherent on its own.²⁷

- (40) a. the women, they read the newspaper
 b. the newspaper, the women read the newspaper

In constituent questions, the added XP includes the DF FOC. Asudeh (2004: 49) uses the term UDF, unbounded dependency function – FOC or TOP – to generalise across unbounded dependencies. Fronted FOC typically occurs with languages such as English, as shown in (41b, d).

- (41) a. Carmen is licking an ice cream-FOC/OBJ in the garden
 b. what-FOC/OBJ is Carmen licking in the garden?
 c. Carmen is licking an ice cream in the garden-FOC/ADJ
 d. where-FOC/ADJ is Carmen licking an icecream?

²⁷ When the topicalised NP is internal to the clause and not in its canonical position (and hence ambiguous), its functional uncertainty (Bresnan 2001: 67) is resolved differently in different languages, e.g., by clitic marking on V in Italian (cf. ch. 3), and case marking on N in Japanese (cf. ch. 4).

And, indeed, it is a well-attested fact in English interlanguage that learners can front their question words at this stage. What they cannot do yet is disrupt the canonical word order frame.

At the next stage, the crucial step forward is the ability to scramble the elements of the canonical word order. What enables this to happen is that the learner can now assign a GF to each of the arguments autonomously, that is, abstracted from the fixed position they occupy in the canonical order frame (e.g., by unequivocally identifying SUBJ through agreement with V). This makes argument functions other than SUBJ sufficiently independent as to receive, themselves, the assignment of TOP or FOC. All this certainly needs the S-procedure to be firmly in place. If no functional assignment markers (such as those provided by morphological resources, cf. § 4.3) were to signal to listeners that the first NP is not SUBJ, they might interpret it as SUBJ, and misunderstand the message. Conversely, the S-procedure being in place does not guarantee that the learner is also able to topicalise OBJ or other argument functions.

In sum, in order to capture the learner's syntactic progress from the simple and strict canonical word order frame to a richer and more flexible one, the Prominence Hypothesis predicts three stages: first, word order is canonical, any other additional information may be placed after it. TOP and SUBJ coincide by default, with minimal (positional) specifications. Second, word order is still canonical, but because an XP is added as TOP or FOC, DFs and GFs are differentiated. Finally, word order may be other than canonical, and GFs are further specified.

Empirical evidence for the Prominence Hypothesis is reported in several studies across different languages and situations, such as Kawaguchi (2005, 2010) for Japanese, Di Biase (2007) and Bettoni & Di Biase (2011) for Italian, Zhang (2007) for Chinese, Yamaguchi (2008) for English, and Itani-Adams (2009) for Japanese-English bilingual acquisition, all of them within the PT framework.

4.2.2. The Lexical Mapping Hypothesis

As we have just seen in dealing with the Prominence Hypothesis, the speakers' assessment of which referent may be salient in the current communicative context will contribute to the assignment of prominence to a particular argument, whose alignment may yield a noncanonical word order construction aiming to achieve a given pragmatic effect. Similar discourse-pragmatic effects may also be achieved by selecting a verbal lemma or construction that requires nondefault mapping between thematic roles and GFs. Importantly, though, topicalisation or focalisation actually preserve the semantic role hierarchy. So, if agent is mapped on SUBJ, topicalisation does not affect this relationship, as we have seen in the Russian sentence in (27) above, which topicalises OBL_{GOAL}. On the other hand, according to LFG Lexical Mapping Theory, the default mapping of thematic role onto GFs may be

altered when, for example, the patient role (rather than the agent) maps onto SUBJ, at the same time leaving canonical word order unaffected (i.e., SUBJ may still be in first position). An important set of verbs marking saliency through non-canonical mappings are the passive verbs, which are present in many languages. Also the so-called exceptional verbs with intrinsic lexical features that require non-canonical mapping as their ‘normal’ (i.e., pragmatically unmarked) realisation are present in many languages. We have already illustrated both an exceptional verb and a passive one in (26) and (27) respectively in § 2.2.

With regards to the learner’s progress, the Lexical Mapping Hypothesis was originally proposed by Pienemann, Di Biase & Kawaguchi (2005: § 3.8) to trace the way learners progress through the staged development of syntax beyond the rigidity of canonicity towards a fuller flexibility of the optional choices allowed by their L2 in assigning GFs to thematic roles. We now formulate an explicit and substantially expanded Lexical Mapping Hypothesis interfacing syntax with both discourse and semantics in (42)²⁸, summarise it in (43), and then comment on it.

- (42) *Lexical Mapping Hypothesis*. Second language acquirers will initially map the highest available role in the thematic hierarchy (e.g., agent, experiencer) onto minimally specified SUBJ/TOP. We call this *default mapping*. Next, they learn to add further arguments mapped onto grammatical functions (GFs) differentiating them from SUBJ (and OBJ, if present). They may also learn some exceptional verbs at this second stage. Finally, they learn to impose their own perspective on events, that is, to direct the listener’s attention to a particular thematic role lower in the hierarchy by promoting it to SUBJ, and defocus the highest role by mapping it onto a GF other than SUBJ, or suppress it altogether. At this last stage learners may add further role information regarding causality, benefit, or adversity. They may also add to their lexicon particular subsets of Vs, such as unaccusatives, as well as further intrinsically exceptional Vs requiring their own mapping schema. We call this *nondefault mapping*.

- (43) *PT: syntactic development based on the Lexical Mapping Hypothesis*

STAGE	CONSTRUCTIONS
NONDEFAULT MAPPING	unaccusatives, passives, causatives, exceptional verb constructions, etc.
DEFAULT MAPPING AND ADDITIONAL ARGUMENTS	agent/experiencer mapped on SUBJ, patient/theme mapped on OBJ, and other members of the a-structure hierarchy, such as goals and locatives, mapped on OBL
DEFAULT MAPPING	agent/experiencer mapped on SUBJ and patient/theme mapped on OBJ
LEMMA ACCESS	single words formulas

28 An earlier formulation of the Lexical Mapping Hypothesis is found in Kawaguchi & Di Biase (2005).

To begin with, soon after the stage of single words and formulas, a default choice for learners is to map the agent or experiencer role as SUBJ (positionally prominent) and the patient or theme, if present, as OBJ. This indicates, minimally, that SUBJ and OBJ are differentiated.

Then, after the initial default mapping stage and before the final nondefault mapping stage – differently from the Pienemann, Di Biase & Kawaguchi (2005) hypothesis, repeated almost unaltered in Pienemann (2011c: 48, fig. 3.13) – we hypothesise an intermediate stage. At this stage, crucially, learners add a further argument role such as a goal/beneficiary, instrument, or locative, mapped onto an OBL GF. Should learners add this further argument without functional marking, confusion could arise as to who does what to whom in the eventuality, especially if there are two animate or human participants as in (44a). But if they add it with functional marking, they are able to differentiate between core and noncore arguments: typically NP vs PP, as in (44b), but also by means of position, as in (44c).

- (44) a. *Romeo gives a rose Juliet
 b. Romeo gives a rose to Juliet
 c. Romeo gives Juliet a rose

At its final stage, the Lexical Mapping Hypothesis is quite broad and undifferentiated, as can be gathered by the wide range of verbs and verbal constructions listed in the highest row in (38), which includes causatives, benefactives, and unaccusatives, among others. These verbs and constructions are arguably among the most language-specific. Within this framework, what unites them, both within a language and across languages, is that the nondefaultness of the mapping of a-structure to f-structure is determined either by the speaker's lexical selection alone or by lexical selection in conjunction with discourse-pragmatic reasons. By contrast the noncanonicity in the last stage of the Prominence Hypothesis is determined exclusively by discourse-pragmatic reasons. What unifies them from a language acquisition point of view is that, until GFs are unequivocally assigned to thematic roles, canonical word order associated with nondefault thematic mapping can mislead the hearer, especially when both participants are animate, as in (45a), where the speaker's intention is to communicate the message in (45b).

- (45) a. *[call your Mother,] she worries you
 b. you worry her / she worries about you

The sequence in which all these exceptional V and nonbasic V forms are learned within the nondefault mapping stage is a widely open empirical question. It may well be, for example, that learners acquire at least some exceptional Vs earlier than

passive Vs, because in the case of exceptional Vs the motivation for the Lexical Mapping Hypothesis is triggered exclusively by their intrinsic lexical nature (features) – that is, each particular V must be learned as an exemplar –, whereas in the case of nonbasic passive V forms learners may find out, eventually, that the passive construction is generalisable across transitive Vs. However, the assumed sequence needs to be tested crosslinguistically.

Empirical evidence for the Lexical Mapping Hypothesis comes mainly from Kawaguchi's study of Japanese L2 (cf. Kawaguchi 2005, 2007, 2009a, 2010), two studies on the acquisition of passives in English L2, namely Keatinge & Kefßler (2009) and Wang (2009), and one by Bettoni, Di Biase & Nuzzo (2009) on the acquisition of postverbal SUBJ in Italian L2. Less robustly supported by crosslinguistic empirical evidence than the Prominence Hypothesis, this Lexical Mapping Hypothesis needs further testing across different languages and communicative situations.

4.3. Interfacing developmental schedules

As we have seen, with regard to morphology PT claims that progress is constrained by the syntactic level of the constituents whose features require unification (cf. § 4.1). This means that learners develop first lexical, then phrasal, and finally interphrasal, morphology – we do not deal here with interclausal morphology because, as mentioned in § 4.1, the complex area of subordination remains a gap in PT, and requires further theoretical elaboration and focused empirical investigation. With regard to syntax, progress is determined by the type of correspondences that map constituents (c-structure) and argument roles (a-structure) onto GFs (cf. § 4.2). This means that learners develop from a rigid canonical word order to more flexible choices, and from default mapping between thematic roles and GFs to nondefault mapping. How then may the parallel schedules we presented separately in (31), (35) and (43) interface? When in 2005 Pienemann, Di Biase & Kawaguchi added their extension to Pienemann's original PT version, the interface between "morphosyntax" in 1998 and syntax in 2005 was often mentioned, but not worked out in any detail, nor has this been treated further by Pienemann (2011a, b, c) more recently. Here we can only state the problem in its broader form, because the details will necessarily be language specific (cf., e.g., the early mention for L2 Italian in Di Biase & Kawaguchi 2002; the treatment of case in L2 Russian and L2 Serbian, and of Spanish Differential Object Marking, chh. 5, 6 and 7 respectively, this volume).

We assume that initially, when the activation of the category procedure distinguishes at least Vs from Ns, the general tendency for learners would be to organise their utterances syntactically by mapping their a-structure roles onto GFs in a default way, and sequencing them in canonical order in c-structure. However, on

the one hand, the category procedure (i.e., differentiation in the lexicon) is a necessary resource for both canonical word order and default mapping. On the other hand, canonical word order and default mapping do not necessarily share the same constraints. As a matter of fact, under pragmatic pressure, learners may not map the highest thematic role available (e.g., the agent) onto the most prominent position or the highest GF, and this provides a good reason for us to abandon the 2005 Unmarked Alignment Hypothesis. For example, under clear pragmatic conditions urging the learner to topicalise the theme argument rather than the agent argument, learners at this early stage may well produce sentences such as (46a), where the SUBJ_{THEME} V OBJ_{AGENT} string would show nondefault mapping with canonical word order – leaving the interlocutor with an ambiguous or uninterpretable meaning, although in this particular example animacy reduces the conflict. (For further examples and discussion on this point, cf. ch. 3 on Italian, § 3.1, this volume).

- (46) a. *the lettuce eat the goat
 b. the goat eats the lettuce

At an intermediate level, in order to progress beyond canonical word order, learners need to differentiate between SUBJ and TOP or FOC. This may be achieved lexically. That is, as long as the lexicon includes temporal or spatial adverbials such as *yesterday* and *in Libya* (cf. (39a) above) to express TOP_{ADJ} in declaratives and question words such as *what* to express FOC in interrogatives, learners may not need to be beyond the category procedure stage. Likewise, progress beyond default mapping may be achieved lexically in many languages by distinguishing between unmarked OBJ and an OBL marked by a preposition. On the other hand, the phrasal procedure would be necessary within the VP in languages which mark GFs by case rather than by PP (cf. the discussion for Russian L2, ch. 5, § 3, this volume). Finally, with regard to the interface between the two syntactic schedules, there seems to be no reason why the addition of an XP before canonical word order should presuppose the ability to add another argument to the verbal frame, or vice versa.

As we have just seen, language specificity plays an important role in the interfaces among the schedules at the intermediate level. This is even more evident at the advanced level. Here the key issue is mature, unequivocal functional assignment. The means – or, more precisely, the mix of means in competition with each other (cf. Bresnan 2001) – that languages deploy to assign GFs and DFs to thematic roles and constituents include, typically, morphological means, but also lexical, syntactic, as well as prosodic means. Prosody is a big player here (cf. Levelt 1989), but this area is beyond the current scope of PT. This variety of means entails that progress beyond the initial stage along the three schedules may vary

substantially from language to language, as the three chapters in part II of this volume will illustrate.

Even within a single language it may be premature to predict implications among the schedules. In English, for example, learners will have to reach the sentence procedure stage in order to produce sentences with either noncanonical word order or nondefault mapping, but with regard to the interfacing between the Prominence Hypothesis and the Lexical Mapping Hypothesis (Kawaguchi & Di Biase 2012), it is empirical evidence that should show the sequence, if any, in which structures such as those in (47b-d) will be learned after (47a).

- | | |
|-------------------------------------|-----------------------------------|
| (47) a. <i>canonical word order</i> | Juliet kisses Romeo |
| b. <i>topicalisation</i> | Romeo Juliet kisses |
| c. <i>passive verb</i> | Romeo was kissed [by Juliet] |
| d. <i>exceptional verb</i> | Romeo received a kiss [by Juliet] |

Suffice it to mention here three issues affecting the interface between the two hypotheses. First, what motivates them? We assume that, with regard to the Prominence Hypothesis, canonical vs noncanonical word order is triggered only by discourse-pragmatic motivation, independently from the feature structure of the V lemma. On the other hand, with regard to the Lexical Mapping Hypothesis, default vs nondefault mapping is triggered either exclusively by the intrinsic lexical nature of V (e.g., in the case of exceptional Vs), or by the lexical features of V in interaction with discourse-pragmatic attribution of TOP or FOC (e.g., in the case of passives). Secondly, Vs requiring ‘only’ intrinsically nondefault mapping may be relatively easy to acquire. The English V *receive* is one of them. Although it requires that the beneficiary, rather than the agent, be mapped onto SUBJ, the beneficiary is the highest role in its <beneficiary, theme> a-structure, and the theme (which is hierarchically lower) is regularly mapped onto OBJ. Moreover, word order is canonically SVO. Conversely, some other Vs are much harder to acquire. The Italian V *piacere* is one of them (Bettoni, Di Biase & Nuzzo 2009). This may be related to the complexity of their lexical features, which must be learned V by V. Exceptional Vs, in fact, involve what Skehan (1998) calls “exemplar-based knowledge”, as opposed to “rule-based knowledge”, because the behaviour of each of them cannot be rule-generated and is not generalisable to other Vs. Since exemplar-based knowledge occupies a larger memory storage than rule-based knowledge, it is less economical for online language production. Thirdly, the alternatives in (47), as well as others not exemplified there, are not all available in all languages, nor do different languages show the same preferences among all available options – as the awkwardness of some of the English ones in (47) imply.

5. Methodological issues

Parts II and III of this volume will empirically test, language-specifically and on a wide typological range of L2s, the broad learner's progress outlined in this chapter within its universal, processability-based framework. Before proceeding, however, we need to mention at least three methodological issues of crucial importance to PT: (a) the type of data required for testing the hypotheses; (b) the type and number of structures which can be diagnostic for a given stage; and (c) the operational criteria used for determining the acquisition of specific structures.

The first issue concerning the type of data needed to test PT hypotheses is the least controversial. PT is a theory of SLA which explains the learner's development on the bases of speech processing procedures – and can thus, in Ellis' (1994: 682) words, be defined as a “transition theory” (i.e., concerned mainly with the dynamic mechanisms responsible for the changes in the development from one system to the next) rather than a “property” theory (i.e., concerned mainly with describing static systems of linguistic knowledge). As such, PT relies primarily on spontaneous speech data, produced online, in order to infer developmental patterns, as already proposed in Pienemann (1998), and – we may add – in most L1 acquisition research, where there might be very little point in trying, for instance, grammaticality judgment methods. Indeed, also linguists interested in knowledge acquisition, whether L1 acquisition or bilingual L1 acquisition, rely heavily on children's performance data to infer knowledge patterns. Thus PT does not use as primary evidence other types of data such as grammaticality judgements or introspective comments used for instance by property theories working within Typological Markedness or Universal Grammar frameworks (cf., e.g., Doughty 1991; Gass 1994; Sorace 2003; White 1989). Such data, usually cross-sectional, is elicited in order to assess the speaker's grammatical knowledge (Loewen 2009), whereas PT is more interested in accounting for performance. It would be also rather difficult for grammaticality judgements (which are comprehension-based) to deal, for instance, with optional DFs or features that rely on speaker-induced choices in production. Sure enough, measures from laboratory data such as reaction time, used also in Pienemann (2002), may be very useful for testing whether some particular elements are, or have become, part of learners' procedural knowledge, but they may not be as useful for looking at developmental patterns. Inferences about development are ideally drawn from oral longitudinal data, although cross-sectional data is also often used, recruiting informants over a range of proficiencies. Again, ideally, cross-sectional data may best be used to confirm developmental patterns inferred from longitudinal data.

PT's primary reliance on spontaneous spoken data does not rule out the use of specifically devised communicative tasks to elicit structures which occur less frequently in natural conversation or emerge only in highly specific contexts, as is

often the case for optional structures which are highly marked and depend on the speaker's particular pragmatic choices. What PT requires is the assurance that the sentences uttered by learners are computed online thanks to their speech-procedural skills, at whatever stage of development they may be. For this reason spoken production, which moves at a rate of between two and three words per second (Levelt 1989) is privileged in PT over written production, which allows for much longer processing time.

The second issue concerns the type and number of structures needed as evidence that a learner has reached a given stage. The answer to the number of structures is fairly straightforward: one is sufficient. In fact, according to PT, learners are deemed to have reached a certain stage when they show their ability to activate the relevant procedures in their L2, and such evidence can be obtained by producing a single type of structure, for example, single words and formulas, canonical order, and so on. The answer to the question of which type of structures the researcher chooses is more delicate because not all structures belonging to a stage will be acquired at the same time, and we may well find that the 'earliest' (i.e., easiest) structure at the following stage is acquired before the 'last' (i.e., most difficult) of the previous stage. This is what Pienemann (1998: 153), following Larsen-Freeman & Long (1991), calls 'scouting' and 'trailing'. The best choice for a diagnostic structure on an untried language should fall on a structure that displays possibly the clearest one-to-one relationship between form and function, or the most representative, or default, structure of a stage in a particular schedule, the one with the most transparent conceptual meaning. Language specificity plays a crucial role here. Let us take morphological development in a language such as Russian as an example. Russian requires number, gender and case agreement between noun and adjective within the NP. So, in order to place a learner at the phrasal procedure stage, any of these three features would do, and there is no theoretical need to limiting or pre-empting the search – provided, of course, that in a fusional language such as Russian the concept of "factorisation" (Pienemann 1998: 159) is used to disentangle the three features. However, the number feature is conceptually more transparent than either gender (which is lexically assigned) or case (which is structurally and/or lexically assigned). Number is also more common typologically. So it would seem to make sense to pick this feature to begin with. If we were to choose the gender feature, for example, we may well find that it is acquired after the emergence of number agreement at the higher sentence procedure stage, which may lead researchers to conclude that PT's prediction is wrong. Charters, Jansen & Dao (2011) is a case in point: they notice that Vietnamese children learning English produce numerical phrasal plural before categorical plural, and claim that this conflicts with PT's hierarchy (i.e., categorical before phrasal procedure). But, before venturing into that claim and inval-

idate PT, they would have to show that when numerical phrasal plural is produced no other categorial marking (e.g., *-ing* for V-like items) has emerged in the L2. It could well be that Vietnamese children may have a further barrier with English plurals but not with categorial marking *tout court*.

All this does not mean, of course, that researchers should focus only on the ‘earlier’ structures of a stage. It is indeed desirable to gain an understanding of the range of structures that may belong to a particular stage, but the bottom line is that the implicationality of the staged development can be tested on just one structure for each stage in a kind of ‘rapid profile’ (cf. Pienemann & Kessler 2012). On the other hand, this cannot be the full story for PT. Indeed some researchers (e.g., Mansouri & Håkansson 2007) working on Arabic and Swedish respectively, found that some structures within a particular stage appear to emerge regularly after others, which prompted them to conceptualise an ‘intrastage’ sequencing. We believe that intrastages are likely characteristic of almost any stage in any language. These may be due to a range of reasons: intrinsic lexical features, such as with exceptional Vs or N classes in particular languages (cf., e.g., Di Biase 2008 specifies extra features for the production of certain plurals due to N class in Italian), or particular extra steps linked to the processing of certain structures (e.g., computing tense and aspect choices in VPs). In this volume we propose to call these ‘extra’ processing barriers within a stage ‘soft barriers’. This means that, once the hard barriers across stages have been crossed, language specific soft barriers within that stage can be identified, and attendant hypotheses may be entertained and tested. In later chapters language-specific patterns may be found to exemplify this phenomenon. For a puzzling example, see in chapter 2 the late acquisition of the so-called ‘regular’ past marker *-ed* in English, which belongs to the categorial stage but whose emergence coincides with much later stages. In other words, as Lardiere (2009) puts it, “(a)ssembling the particular lexical items of a second language (L2) requires that the learner reconfigure features from the way these are represented in the first language (L1) into new formal configurations”. Within PT, we would wish to know the relative position in the sequence of any L2 grammatical structure and, possibly, pinpoint the feature (bundle) that may cause them to occur when they do.

The third issue is more controversial and concerns the definition of acquisition criteria that allow us to determine operationally whether a given structure has been acquired or not. Since its inception, PT has used the emergence criterion, that is, “the point in time at which certain skills have, in principle, been attained or at which certain operations can, in principle, be carried out” (Pienemann 1998: 138). For a critical review of Pienemann’s theoretical and empirical construct and a thorough discussion of how the emergence criterion can be best formulated and operationalised with regard to morphology, we refer to Pallotti (2007). Suffice it to say here that the authors of the following chap-

ters in this volume all follow the emergence criterion. However, generally following Pienemann (1998: 133), their way of operationalising it for morphology differs from the one used in syntax. On the one hand, the obligatory nature of morphological structures allows us to determine the contexts in which a particular structure occurs, and hence to observe and compute their distribution over the corpus, that is, whether they are supplied or not supplied in obligatory contexts, and/or produced where they are not required. On the other hand, this is not possible for syntactic structures, whether default or optional, since they mostly depend on pragmatic choices. Hence only the positive evidence of the number of their valid instances can be counted.

In general terms, we may say here that, for morphology, we accept as evidence of emergence the production of two occurrences of a structure, provided there is formal and lexical variation among the two. This means both morphological variation (e.g., *go* vs *goes*, a kind of structural minimal pair) and lexical variation (*goes* vs *eats*, a sort of lexical minimal pair), as Pienemann has always stated – although Keßler & Pienemann (2011: 95) do not commit themselves to any number, simply saying in their example that “various” Vs are needed for lexical variation, and “different morphological forms” are needed for “some of these verbs”. This principle of structural and lexical variation is designed to flush out exclusively formulaic use. When agreement structures involve more words, a greater number of occurrences may be in order, but the design principle remains the same: minimal evidence of structural and lexical variation is required. For syntax, following a long tradition from ZISA (Meisel, Clahsen & Pienemann 1981) to PT, as in Pienemann, Di Biase & Kawaguchi (2005) we accept one example as evidence of the emergence of a structure. For further details, we refer to the individual studies reported in the chapters to follow.

Let us then finally describe in (48) how the implicational scales in the subsequent chapters of this volume organise the empirical data for testing PT’s hypotheses.

- (48) a. We follow Pienemann’s long tradition of indicating progress vertically by placing the earlier structures at the bottom row, and then proceeding upwards to the most advanced ones at the top. Similarly, the learners’ progress in time is shown by proceeding left to right, that is, from *t*1 to *t*_n in longitudinal studies, and from the least advanced learner to the most advanced one in cross-sectional studies.
- b. The lighter horizontal lines mark the divisions between the stages, thus grouping together all the structures within each stage. We call these interstage divisions *hard barriers* for the learner to negotiate.
- c. The numbers of occurrences for morphological structures are entered in the tables with a plus (+) sign if the structures are supplied in obligatory contexts; with a

minus (-) sign if not supplied in obligatory context; and with the greater-than (>) sign if supplied in a context that does not require it; an empty cell marks no context for a structure. As for syntactic structures, their optional nature makes it futile to establish obligatory contexts. So the relevant cells only mark their presence, and show either clear numbers or empty cells, unless a structure is attempted but produced in an interlanguage form, in which case the number is preceded by a minus (-) sign. For example, in *the green fish eaten by red fish*, a passive structure is attempted, but the auxiliary is not supplied.

- d. In order to highlight the implicationality of the learners' progress, thicker vertical lines mark at which time (longitudinally) or by which learner (cross-sectionally) a particular *stage* has been reached – in other words, whether the acquisition criteria are not satisfied (to the left), or are satisfied (to the right). These vertical lines are drawn for the whole stage, that is, regardless of whether only one or more of the structures belonging to that stage have emerged for that stage. Because – as we have just mentioned above – one contrastive token for syntactic structure (or two lexically contrastive tokens for morphology) is sufficient evidence for declaring that a *stage* has been reached, it may well be that negative evidence for a different structure belonging to the same stage appears to the right of the vertical line. When this is the case, a zigzag vertical line marks at which time (longitudinally) or by which learner (cross-sectionally) a particular *structure* has been acquired. We have called these intrastage divisions *soft barriers*, which the learner has to negotiate further within the stage. To illustrate this, we offer the following table:

STAGE	STRUCTURE	T1	T2	T3	T4
4	E			-1	+2
3	D	-1	-2	+2	+3
2	C		-2	-1	+3>1
	B	-1	+2-1		+4
1	A	+2	+3-2>2	+4-1	+5

Here stage 1 is reached at t1, stage 2 at t2, stage 3 at t3, and stage 4 at t4. At t2, however, only one (structure B) of two structures in the same stage is acquired. The other (structure C) is acquired at t4, that is, after the soft barrier is overcome, and after structure D of stage 3. In English L2, for example, the *-ed* morpheme for regular past, which belongs to the category procedure stage, is usually acquired only when higher stages are reached (cf. ch. 2).

- e. Where appropriate, a scalability statistical test (cf. Hatch & Lazaraton 1991: 204-216) is used to show the existence of hierarchical relationships among structures.

6. Conclusion

In 1998 Pienemann defined his theory as 'lexicalist'. Yet, PT's main concern was then limited to the procedural activities of Levelt's formulator and did not deal with the lexicon in sufficient detail except for its diacritic features. Further, 1998 PT was powerful in explaining obligatory morphological variation, but did not tackle syntactic issues of marked word orders or mapping. Thus it could not explain why structures such as topicalisation or passives emerge in interlanguage much later than their canonical SVO/SOV or active counterparts respectively. Pienemann, Di Biase & Kawaguchi's 2005 extension brought into the scope of PT structural consequences triggered by discourse-pragmatic and lexical requirements, which inevitably interface with syntax. In this chapter, we have proposed an updated PT formalism, and made more explicit connections between PT and Bock & Levelt (1994), as well as Levelt, Roelof & Meyer's (1999) Theory of Lexical Access on the one hand, and Bresnan's (2001) and other LFG scholars (e.g., Dalrymple, Falk) formalisation of grammaticalised DFs and Lexical Mapping Theory on the other. We have thus coherently broadened the scope of PT to include FOC in addition to TOP in the syntactic schedule, as well as recast the syntactic hypotheses, and generally tidied up the schedules resulting from the Prominence Hypothesis and the Lexical Mapping Hypothesis, both of which could not be coherently understood, nor formalised in fact, without the later psycholinguistic and grammar-theoretical works just mentioned.

We have also proposed a more explicit interface between morphology and syntax. The acquisition of obligatory morphology is explained in Pienemann (1998) by the sequential activation of Levelt's category, phrasal and inter-phrasal procedures modelled in LFG by the formalisation of feature unification. Ultimately, this also explains the later acquisition of structural alternatives, because their noncanonicity is marked by some morphological or other (nonsyntactic) means in many languages, including those with scant morphology, such as Chinese, for instance, where a small number of particle markers become crucial as indicators of mature functional assignment.²⁹ Indeed, it would be interesting for research to look at the applicability of our Prominence Hypothesis to so-called topic-prominent languages (Li & Thompson 1981) such as Chinese. Compared to their canonical or default counterparts, the later acquisition of the noncanonical or nondefault structures accounted for by the Prominence Hypothesis and the Lexical Mapping

²⁹ Prosodic means are of course important. However, in order to deal with them in a formal way, PT researchers will require further work within LFG and Prosodic Phonology (Nespor & Vogel 2010).

Hypothesis does not by itself determine higher stages in interlanguage; noncanonicity and nondefaultness however are dependent on morphological resources, in the sense that their fluent deployment implies the automatisisation of morphological components.³⁰ So Pienemann's (1998) PT remains, in any case, foundational.

The crucial question in order to move from 'descriptive adequacy' to 'explanatory adequacy' is what motivates alternative syntactic structures. According to Levelt's Model, the preverbal output of the conceptualiser guides the order and type of lexical choices. When learners must compute them online, including their phonological and prosodic shapes, along with discourse-pragmatic information and/or lemmas with richer feature structures, and integrate them all within an executable frame for oral production, these extra processing computations add to the cognitive cost for the learner (and natives, though minimally, cf. Weyerts et al. 2002). However, the cost of effective communication is reduced by automatisisation, and if one can handle the phonology and syntax automatically, then more attention can be paid to processing semantic, pragmatic, and sociolinguistic levels of communication (cf. Segalowitz 2003). This is a processing area that cries out for integration in a theory which has processability at its heart.

Part II of the volume will now illustrate how our view of PT interacts with language specificity.

30 This point is not uncontroversial and invites focused research.

PART II

The developmental path across languages

Camilla Bettoni* and Bruno Di Biase**

*University of Verona, **University of Western Sydney

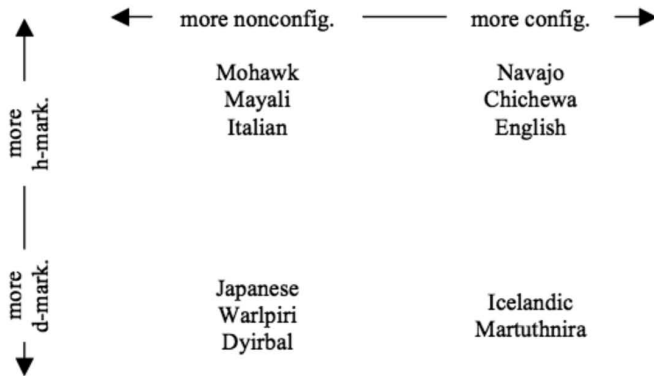
After focusing on the universality of the theory and on the integration of its 1998 and 2005 strands into a more coherent whole in part I, this volume, part II draws the consequences of these two foci, and reconceptualises the staging of L2 development with reference to three typologically distant languages covering a good chunk of typological space between them: English, a configurational language; Italian, a null-SUBJ head-marking language; and Japanese, a zero-anaphora, dependent-marking language. The latter two languages are all placed towards the less configurational end of the continuum, as shown in (1) below.

The universality of PT “universal schedules” is based on speech processing procedures, which are cognitive, and hence universal. By that, however, we do not mean that every language will have the same developmental schedules. Rather, we mean that the universal schedules can only be interpreted in a language specific way. Thus, every language has its own schedules reflecting its own typology. This is why part II of the volume describes the development of three typologically different languages.

There are two principal sources of language specificity that the learner must acquire (aside from phonological considerations): the lexicon and c-structure. These are linked via f-structure, which is largely universal, but expressed in a language-specific lexicon and aligned according to language-specific constraints interfacing with discourse-pragmatic preferences. In this regard, there are two important typological distinctions – or rather continuums, because natural languages may freely mix their modes of organisation (Bresnan 2001: 132). The first continuum – as we have already seen with the two extreme examples of English and Warlpiri (cf. ch. 1, § 2.2) – is configurationality, which distinguishes between languages expressing GFs (principally the relationship between the verb and its arguments) by position, and those expressing them by morphology. The second important typological continuum relevant to our volume distinguishes between languages marking the relation between the constituents and the head morphologically on the head (such as Italian, and to a lesser extent English), or on the dependent (such as Japanese). This characterization as head-marking or dependent-marking, first introduced by Nichols (1986) for any kind of phrase structure, indicates for

us mainly whether GFs are marked inflectionally on the head element (typically the V, or the predicate) or on the dependent element (typically the nominal arguments). For example, a language is head-marking if it overtly marks the SUBJ function in a clause by means of the agreement of V with its SUBJ; on the other hand, a language is dependent-marking if it marks the NP argument by case-feature. Some languages may use both agreement and case marking (e.g., Serbian and Latin), others hardly any (e.g., Chinese). In (1) below we have added the three languages treated in this part of our volume to the schema introduced by Nordlinger (1998).

(1) *Basic typology of expressing grammatical relations (after Nordlinger 1998: 49)*



With regard to configurationality, represented on the horizontal continuum in (1), we have already shown in chapter 1, (23), how a highly configurational language like English uses hierarchical phrase structure to encode GFs such as SUBJ and OBJ. English in fact is one of those languages where OBJ belongs under VP and is strongly related to V, and ADJ may not be freely interposed between V and OBJ (unlike Italian or Spanish). On the other hand, SUBJ is outside VP and precedes V. English SVO word order is fixed, to the extent that, if the NPs before and after V are swapped, the meaning of the clause changes, as in (2).

- (2) a. Jane hits Tarzan
b. Tarzan hits Jane

At the other end of the configurationality continuum, as we have shown in chapter 1, (24), Warlpiri uses morphological case marking on NPs, rather than syntactic phrases, to encode GFs. This type of marking allows for a highly flexible word order, though a positional point of reference remains even in radically configura-

tional languages, like Warlpiri where, for instance, AUX must be in second position (cf. Asudeh & Toivonen 2010), thus retaining a certain positional organisational principle.

Like Warlpiri, Italian and Japanese are also nonconfigurational languages, although less radically so, in so far as they do exhibit a canonical word order (SVO in Italian and SOV in Japanese), and neither allows elements belonging in the same NP to be easily separated. Both these languages allow for some flexibility in word order. However, they differ from each other because they represent a case of head-marking and dependent-marking languages respectively. We illustrate this difference by looking at morphological encoding of the two core GFs SUBJ and OBJ by means of agreement marking on V in Italian, the more head-marking language, and of case marking on NPs in Japanese, the more dependent-marking. For example, in the two Italian sentences in (3), word orders are SVO and OVS; yet their referential meaning is the same. This is so because, when OBJ topicalisation disrupts canonical word order, the functions of both NP_{SUBJ} and NP_{OBJ} are identified morphologically by two inflections of V: one, which marks SUBJ, is identified by the V morpheme *-a*, agreeing with postverbal SUBJ; the other, which marks OBJ, is identified by *lo*, the ACC clitic marker coreferential with preverbal TOP (for further details on OBJ topicalisation in Italian L2, cf. chh. 3 and 8, this volume).

- (3) a. Desdemona picchia Otello
 Desdemona-3.SG hit-3.SG Otello
 [Desdemona hits Otello]
- b. Otello lo picchia Desdemona
 Otello-3.MASC.SG he-3.ACC.MASC.SG hit-3.SG Desdemona-3.SG
 [Desdemona Otello hits]

Likewise, in the two Japanese sentences in (4) from Kawaguchi (2008: 96), word orders are SOV and OSV respectively; yet their propositional meaning does not change. However, unlike in Italian, this is so because, irrespective of their positions, the function of NP_{SUBJ} is identified morphologically by the case-marking *-ga* for NOM, and the function of NP_{OBJ} by the case-marking *-o* for ACC.

- (4) a. Mari-ga Takashi-o nagutta
 Mari-NOM Takashi-ACC hit-PAST
 [Mari hit Takashi]
- b. Takashi-o Mari-ga nagutta
 Takashi-ACC Mari-NOM hit-PAST
 [Mari hit Takashi]

In sum, different languages encode GFs by different means. English, a configurational language, does it mainly through configurationality, even though it uses some vestiges of head-marking morphology (e.g., 3rd person singular is marked on V in the present tense). Italian and Japanese, both nonconfigurational languages, overtly mark GFs mainly through morphology, and reserve positional options for DFs. PT claims that the learner's morphological and syntactic development can be predicted by

- interpreting the different means by which a target language specifies its grammatical information, representable by an LFG description; and
- identifying the procedural skills required for a particular linguistic operation, as indicated by Levelt's Model.

The developmental hypotheses for English, Italian and Japanese discussed in the following three chapters are not entirely new in themselves, but to a large extent their illustration here is. The changes introduced are consistent with our presentation of PT in part I: they are not mere terminological formalities, but – as we have already mentioned – substantial innovations derived partly from coherently adopting relevant advances in PT's two source disciplines, and partly from our own and the authors' contribution to theory construction and new interpretations of the results in L2 description work. The language-specific developmental schedules presented have been tested to a large extent, albeit some with more robust empirical evidence than others. Where evidence is still scant, it will be indicated in order to identify gaps and promote further work.

Theoretical progress specific to each of the three languages, within a broad roadmap of how they develop in learners, concerns mainly the following areas: a revisitation of the morphological schedules focusing on some neglected areas (e.g., the role and position of the VP procedure), the treatment of questions, both polar and constituent, in English L2 (cf. ch. 2); the identification of soft barriers, or steps, within the stage boundaries, and their explanations for English L2 and Italian L2 (cf. chh. 2 and 3 respectively); and a focus on the Lexical Mapping Hypothesis for Japanese L2 (cf. ch. 4).

2

The development of English as a second language

Bruno Di Biase*, Satomi Kawaguchi*
and Yumiko Yamaguchi**

*University of Western Sydney, **Tokai University

1. Introduction

Many researchers have contributed to understanding the development of English (cf. Finegan 2011 for an overview of this language from a typological point of view), possibly the most studied L2, both outside PT, as can be gathered from well known introductions to the field (e.g., Doughty & Long 2003; Ellis 2008; Larsen-Freeman & Long 1991; Gass & Selinker 2001; Ortega 2009), and within PT. The history of English within PT starts with its precursors (e.g., Johnston 1985a, 1997; Pienemann & Johnston 1987; Pienemann, Johnston & Brindley 1988; Pienemann & Mackey 1993), who found, by and large, morphological sequences similar to those found in early SLA research (e.g., Dulay & Burt 1974; Bailey, Madden & Krashen 1974). However, whereas Krashen (1977), for example, gave no explanation for the sequences except for postulating that the order is ‘natural’, PT (i.e., Pienemann 1998) – through ZISA first, and later Levelt and LFG (cf. ch. 1) – added speech processing and principled staging to the ‘nature’ of morphological sequences, and decidedly also grasped the syntactic nettle. All this produced testable (i.e., falsifiable) developmental hypotheses, as more recent work amply demonstrates for English (e.g., Sakai 2008; Dyson 2009; Zhang & Widyastuti 2010; Charters, Jansen & Dao 2011; Kawaguchi 2013; Pienemann, Keßler & Lenzing 2013; Yamaguchi & Kawaguchi 2014).

We will not review here the variously detailed descriptions of L2 English already offered by these and other authors. We will, instead, ground the theoretical changes proposed in chapter 1, this volume, by tracing the morphological and syntactic development of this configurational language using a detailed, longitudinal, case study. For the development of morphology, we will note three areas where our treatment of L2 English differs most from previous PT ones. First, we re-examine the categorial stage with reference to four structures, that is, *-ing* as a verbal marker, past *-ed*, plural *-s*, and possessive *'s*, a previously neglected area. We discuss possessive *'s* in some detail together with possessive determiners and pronouns, but we exclude both the latter from the PT schedule for English. Second,

we place the unification of AUXs (*be*, *have* and modals) with lexical V within the VP in the phrasal stage, unlike Pienemann who since 1998 deals with English VP only in the context of questions. Third, we adopt for English the concept of ‘soft barriers’ (cf. ch. 1, § 5) within stage boundaries to deal with intrastage sequences, such as the very late acquisition of *-ed* within the category stage.

For the development of syntax, we first explore the proposed Prominence Hypothesis, and treat declaratives separately from questions (both polar and content). Then, we explore the Lexical Mapping Hypothesis and test, on English for the first time, the transitional step between default and nondefault mapping proposed in chapter 1, this volume. We attempt to do all this by thoroughly re-analysing the database from Yamaguchi’s (2010, 2013) two-year longitudinal study of a Japanese L1 primary school-aged child learning English as an L2 naturalistically in Australia.

To make sense of our changes we must refer to Pienemann’s own presentations of the PT hierarchy for L2 English. Among his several schematic versions, substantially unchanged since 1998 (cf. Pienemann 2011a, b, tab. 4.7), we refer to Pienemann’s (2005: 24), reproduced in (1).

(1) Pienemann’s (2005: 24) hierarchy as applied to English

PROCESSING PROCEDURE	L2 PROCESS	MORPHOLOGY	SYNTAX
6 sub. clause procedure	main and sub. Clause		Cancel INV
5 S-procedure	interphrasal information	SV agreement (=3sg -s)	Do2nd, Aux2nd
4 VP-procedure	interphrasal information	tense agreement	Y/N inversion, copula invers.
3 phrasal procedure	phrasal information	NP agr Neg+V	ADV, Do-Front, Topi
2 category procedure	lexical morphology	plural, possessive pro	canonical order
1 word/lemma	‘words’	invariant forms	single constituent

Our main point of contention here concerns Pienemann’s apparent assumption that morphological information and syntactic arrangements depend on the same processing procedure. In fact his presentation suggests that the same processing procedure (column 1) and the same information-exchange process (column 2) apply for both morphology and syntax. So, in stage 4, for instance, the reader would be justified in assuming that ‘tense agreement’ in the morphology column calls for an interphrasal procedure as much as ‘Y/N inversion’ or ‘copula inversion’ in the corresponding column for syntax. This assumption may turn out to be correct, but it clouds the issues under observation, because the morphological ‘tense agreement’ structure refers to V constructions and does not involve the position of SUBJ on the syntactic alignment, and may hence not require the same procedure as Y/N or copula inversion, which does involve the position of SUBJ. In fact some contradiction arises in (1) as to the nature of stage 4: does the VP-procedure actually belong to phrasal procedure or interphrasal procedure? The morphology column would make it ‘phrasal’, whereas the syntax column would make it ‘inter-

phrasal' because it involves the noncanonical position of SUBJ with respect to auxiliaries or copula. In line with chapter 1, this volume, we adopt an analytical approach that looks at morphological development separately from syntactic development, which helps solve this apparent contradiction. Our study points to a clear answer: the VP-procedure belongs to the phrasal procedure in the morphological schedule, as we will see in § 2.

In any case progress in one grammatical area does not guarantee progress in another. As a matter of fact, recent PT longitudinal studies of English L2 (e.g., Yamaguchi 2010, 2013) and of bilingual L1 acquisition involving English and Chinese (Qi & Di Biase 2005; Qi 2011), and English and Japanese (Itani-Adams 2009) report faster growth in syntax than morphology, sometimes even by two PT stages. Bonilla (2014) also found syntax to emerge earlier than morphology in all of her Spanish L2 learners. These are the sorts of reasons that prompt us to present the implicational hierarchy for morphology and its distribution over syntactic levels (cf. § 2) separately from that of syntax (cf. § 3), as Bettoni and Di Biase argue in chapter 1, § 3. This does not exclude that phenomena across morphology and syntax may appear at the same time. We are simply making the point that, as a matter of analytical procedure, it is desirable to separate phenomena that reflect different explanatory principles, leaving their temporal progression to a separate step in the analysis.

A further concern with Pienemann's presentations of L2 English, including recent work (2011a, 2011b: 51-56, tabs 4.1-4.2, 4.4), relates to the fact that morphological development relies on cross-sectional data – chiefly Johnston's (1985a) SAMPLE Report for adult ESL data, and Pienemann & Mackey (1993) for adolescent ESL data. We wish to add here significant support for PT's hierarchy for morphology with genuinely developmental (i.e., longitudinal) data (cf. § 2).

Our proposed changes to Pienemann's English schedules are supported by the evidence afforded by the longitudinal corpus of Yamaguchi's PhD dissertation and subsequent publications (2010, 2013). This data, like most longitudinal data, belongs to one single (child) learner, and provides findings which are hardly generalisable. Yet they are remarkably compatible with those provided by the cross-sectional data from both Johnston's (1985a) adult learners and Pienemann & Mackey's (1993) adolescent learners. What Yamaguchi's longitudinal data contributes, beyond what cross-sectional data possibly can, is to provide critical evidence on the controversial status of some stages (as we have just seen) or particular morphemes, such as plural *-s*, which straddles uncomfortably between categorical and phrasal stages, as well as an idea of the status and temporal sequence of morphemes acquired not only across stages but also within a stage, as we shall see.

The empirical data at hand is produced by a Japanese child, codenamed Kumi, who had been learning English in a naturalistic environment in Australia since her arrival at age 5;7;15. Kumi was first recorded four weeks after arrival (t1),

and then every two weeks for the first 2 months (t2-t5), every two months for the remainder of the first year (t6-t11), and every three months in the second year (t12-t14). Our new analysis of the 14 sessions's data, presented in the next section, covers approximately two years – a 100 week period, up to age 7;8;15. In these sessions various tasks were performed in the child's home with native or near-native speakers of English, including adults and children, who had a close social relationship with her. In order to maintain consistency of data sets across interview sessions, similar tasks were used for each session, including communicative tasks such as 'story telling' tasks, riddles and 'spot the differences' tasks, to elicit questions and declarative sentences. The 'pictures without words' storybook was *Frog where are you?* (Mayer 1969), often used in linguistic research with children, notably by Berman & Slobin (1994). In Yamaguchi's (2010) study this particular task was used in four occasions over the 100 weeks of data collection. Other tasks (e.g., the spot-the-differences task) were also used on more than one occasion, but the stimuli were different each time. The collected data consists of 2,151 turns by the informant with a total of 11,951 word tokens and 3,625 word types (excluding *uhms*, *ers*, and other backchannelling), ranging from a minimum of 176 English words per session at t1 to 1,783 at t14.

2. Morphological development

In (2) we present PT's original and well known universal hypotheses¹ for morphological development applied to L2 English, with morphological structures distributed hierarchically, and implicationally, over phrasal, interphrasal and interclausal levels of information exchange (cf. the universal hierarchy in (31), § 4.1, ch. 1, this volume). Then we follow the learner's progress stepwise, noting the raw occurrences in (3a), and their calculated scalability in (3b). However, within this basic arrangement, in addition to our focus here on the morphological domain to the exclusion of questions and other syntactic structures, we introduce several changes. First, possessive determiners and pronouns do not appear in our hypothesis schedule because they belong to the lexicon, that is, they are not generated by morphological rules. Second, in (2) the order of structures within stages follows the temporal order in which they emerge, when they do emerge, from Kumi's longitudinal corpus shown in (3). So perhaps the most obvious difference between Pienemann's schedules and ours concerns not so much the choice of structures but

1 Needless to say, 'universal hierarchy' in PT never refers to linguistic structures per se, but to their processing, hence the hierarchy may only be expressed linguistically when it is applied to a specific language, English in this case.

rather the addition of language-specific intrastage sequences within stages that have a range of exponents (cf. the notions of ‘hard’ and ‘soft’ barriers introduced in ch. 1, § 5, this volume, and illustrated empirically also for Italian, ch. 3, § 2.2, and for Japanese, ch. 4, § 2.2). For instance, within the category procedure stage, we order the English morphemes as follows: *-ing* emerges before possessive *'s*, which in turn precede plural *-s*, leaving past *-ed* last.

(2) *Developmental stages hypothesised for L2 English morphology (after Pienemann 1998, 2005, 2011a, b)*

PROCESSING PROCEDURE		STRUCTURE	EXAMPLE
S-BAR PROCEDURE		e.g., subjunctive marking in subordination	<i>I suggest he eat less</i> <i>It's time you left</i>
SENTENCE PROCEDURE		SV agreement: 3 rd pers sg <i>-s</i>	<i>Peter loves rice</i>
PHRASAL PROCEDURE	NP PROCEDURE	phrasal plural marking	<i>these girls</i> <i>many dogs</i> <i>three black cats</i>
	VP PROCEDURE	AUX + V: <i>have + V-ed</i> MOD + V <i>be + V-ing</i>	<i>they have jumped</i> <i>you can go</i> <i>I am going</i>
	CATEGORY PROCEDURE	past <i>-ed</i> plural <i>-s</i> possessive <i>'s</i> verb <i>-ing</i>	<i>Mary jumped</i> <i>my brothers working</i> <i>Mary's car</i> <i>he eating</i>
LEMMA ACCESS		single words formulas	<i>station here</i> <i>my name is Pim</i>

(3a) *Kumi's morphological development*

STAGE	STRUCTURE	T1 (4)	T2 (6)	T3 (8)	T4 (10)	T5 (12)	T6 (20)	T7 (28)	T8 (36)	T9 (44)	T10 (52)	T11 (64)	T12 (76)	T13 (88)	T14 (100)
S-PROCED.	3 rd pers sg <i>-s</i>	>2	-2		-1	-1	+1-1	+1-1	-2	+8-1>1	+3	+2-1	+2	+4-1	+35-7
NP PROCED.	quantifier + pl <i>-s</i>		-4	-5		+1-1		+2-2	+1-2	+4-5	+10-8	+4-1	+7-1	+6-3	+13
VP PROCED.	<i>can't/can/will/couldn't V</i>		+1-1	+1-1	+2			+1-1		+4	+7	+7	+6	+4	+13
	<i>be V-ing</i>	-2	+1-1	+1-1	+2-1	+8-2	+5	+11	+10	+7	+5	+4	+6	+6	+7
CATEGORY PROCED.	past <i>-ed</i>		+1-7	-2	-5	-3	-2	+1-3	-2	+5-4>3	+4-1>1	+20>2	+5-2	+19-1>1	+21
	plural <i>-s</i>	-4	+5>1	+1-6>1	+2-3	+2-4	+6-2>7	+2-4	+2-2	+8-6	+9-2	+5-2	+9-1	+9-6	+29-3
	possessive <i>'s</i>		+7			+1>1	+1>1	+3			+5	+2		+3>1	+5>1
	V-ing	+1	+5	+5	+8		+2	+7		+5	+1	+3	+1	+1	+5

in brackets time of session in weeks after arrival

+ = supplied in obligatory context; - = not supplied in obligatory context; > = supplied in wrong context; empty cell = no context

(3b) Scalability matrix of Kumi's morphological data²

		t1	t2	t3	t4	t5	t6	t7	t8	t9	t10	t11	t12	t13	t14
S-PROCED.	3 rd pers sg -s	-	-		-	-	-	-	-	+	+	+	+	+	+
NP PROCED.	quantifier + pl -s		-	-		⊖		+	+	+	+	+	+	+	+
	can't/can/will/couldn't V		-	-	+			+		+	+	+	+	+	+
VP PROCED.	be V-ing	-	-	-	+	+	+	+	+	+	+	+	+	+	+
	past -ed		⊖	⊖	⊖	⊖	⊖	⊖	⊖	+	+	+	+	+	+
CATEGORY	plural -s	-	⊖	⊖	+	+	+	+	+	+	+	+	+	+	+
PROCED.	possessive 's		+			+	+	+			+	+		+	+
	V-ing	-	+	+	+		+	+		+	+	+	+	+	+

At the first lemma access stage, the L2 lexicon contains hardly any grammatical annotations and so learners can activate only the lemma access procedure, and produce morphologically invariant forms. Kumi already produces a narrative of 176 words at t1, hence she is already beyond the one-word/formulaic stage. For this reason, the lemma access stage is not shown in (3a) and is also, naturally, discounted from the scalability calculations in (3b).

Category procedure stage. At the category procedure stage, learners are hypothesised to start annotating lemmas with their characteristic diacritics. These annotations lead the charge towards language-specific differentiation as they are the basic building blocks of the L2-specific grammatical system, starting from the formal differentiation of broad categories such as V-like words from other lexical items. This categorial stage is the most crowded stage for English: PT hypothesises no less than four separate morphological structures to emerge, two characterizing nominal items and two verbal ones, each annotated with its own nature and function, and likely to be used by the learner in ways which may be different from native use. These structures are not required to emerge *all* at the same time (and they don't, as we shall see) nor are they *all* meant to emerge before any morphological structures of the next stage. As far as PT's predictions are concerned, it is enough for any one structure belonging to a stage to emerge before, or coincidentally with, any one structure hypothesised at the next stage. If the data shows that any structure hypothesised to be at a higher stage appear before any structure from a lower stage then the theory is falsified.

2 In a longitudinal study, once the emergence criterion is met in a session, insufficient evidence in subsequent sessions is irrelevant. Hence, some cells with '+1' in (3a) are interpreted as '+' in (3b).

So, what unites the four morphemes within the categorial stage, despite their differences? And in what way are they different from morphemes hypothesised at higher stages? The PT answer³ lies in whether or not they require grammatical unification with other lemmas in the phrase or sentence structure (Pienemann 1998: 97; and (12), § 2.1, ch. 1, this volume). The four morphemes at the English categorial stage do not require grammatical unification with other constituents of the phrase or sentence in which they may appear, so they belong to the categorial stage. We will now proceed to discuss them following their order of emergence as shown in (2). This order is not a necessary part of the PT hypothesis but a matter of observation in Kumi's data. However, we do attempt to explain, in a principled way, early appearance of some morphemes in this learner, such as *-ing* and possessive *'s*, as well as the late appearance of other morphemes, such as *-ed*.

The earliest lemma annotation in English is *-ing*, the key for the L2 learner to appropriate the minimal cognitive resources necessary to differentiate lexical items explicitly: one category will be marked with *-ing*, and others will not, as in (4a-b) where V-like items are marked with *-ing* and N-like and other items are not.

- (4) a. t1 and deer running
- b. t3 mm. girl looking in the hole [...] and girl climbing on rock
- c. t6 and my dad love er.. using this for the drinking
- d. t9 he just kept walking
- e. t13 he saw the man bringing it

This is consistent with Johnston's (1994: 15) observation that "the first function of the *-ing* marker is to enable the learner to build up a prototype grammatical category". Verbal aspect or tense is not crucial yet, so there is little point in calling the *-ing* morpheme a PROG(ressive) aspect marker at this stage. In fact, learners will typically begin by marking V-like items categorially with the *-ing* morpheme (*go* vs *going*) regardless of tense and/or aspect, and without the native auxiliaries. Hence this *go* vs *going* alternation may capture differences in function which, initially, do not correspond to the use native speakers make of these two forms. English *-ing* is in fact an example of a form covering many functions and entertaining relationships within different sorts of structures (e.g., with AUXs for aspect, but also as complement, ADJ and so on). Learners will continue to develop these associations over time, eventually achieving quite sophisticated use, as (4c-e) testify. In the case

3 Morphemes within one and the same categorial stage in PT are notoriously placed at the extremes of the time range for instance in Krashen's natural hypothesis based on the 'morpheme studies': *-ing* is at the bottom, plural *-s* in the middle, and regular past *-ed* at the top of the temporal range.

of Kumi, *-ing* is confirmed as the earliest morpheme to appear with five occurrences on different Vs already at t2, that is six weeks after her arrival in Australia. Other diacritics such as those for marking possession and plurality on Ns or tense on Vs will soon follow, but the important first step has been taken, and the categorical stage can be said to have emerged.

To continue with the order in which morphology emerges in Kumi at this category procedure stage, we now deal with the possessive marker *'s*, and look at it – as Pienemann and Johnston do – in comparison with possessive determiners, which are its lexical counterpart (cf. the table in (5) below).

Before looking at Kumi's data, let us however consider for a moment the position of possessive *'s* and possessive pronouns in Pienemann's English schedules. Way back, in his ground-breaking SAMPLE report (1985a), Johnston places possessive pronouns at his stage 2 (now category procedure stage) and possessive *'s* at his stage 4 (now phrasal procedure stage), presenting substantial data for possessive pronouns (pp. 340, 343) but lamenting the paucity of occurrences for possessive *'s* (19 instances sparsely and unevenly distributed over 12 of his 48 interviews), as well as the lack of 'developmental patterning' and the 'equivocal' nature of the evidence for this form in his cross-sectional study (p. 256). Note that Johnston's learners were of Polish and Vietnamese background, in equal numbers. As we will see later, our Japanese background child informant presents abundant data for this structure. Pienemann & Johnston (1987) and Johnston (2000) keep possessive *'s* at their stage 4.⁴ From 1998 onwards, Pienemann places possessive *'s* at the categorical stage whenever it is mentioned. However, this structure is not mentioned consistently: for example, in Pienemann (2011b) it first appears in the hypotheses (tab. 4.1, p. 51), then it disappears from the tables reporting the adult SAMPLE data (tab. 4.2, p. 52), and the child ESL data (tab. 4.4, p. 56). Possessive pronouns are another area for which the treatment is rather erratic: in his 'possessive pronouns' tables, Johnston (1985) conflates both possessive determiners (or adjectives) such as *my* and *her*, and possessive pronouns proper such as *mine*, *hers*. This may be the meaning of 'possessive pronouns' in subsequent works such as Pienemann & Johnston (1987) and Pienemann (2011a, b). In the latter, possessive pronouns are ignored in tables 4.1–4.2 but appear in table 4.4, where however only a + sign (with no figures) is given.

We prefer to keep the term possessive determiner for adjectives alone, which in Kumi's data emerge earlier and far more forcefully than pronouns, as (5) shows. Also, as proposed earlier, we exclude both possessive determiners and pronouns

4 Krashen's (1977) 'natural order' derived from the accuracy-based morpheme studies places possessive *'s* in the same 'box' as 3rd PER SG *-s* in Vs. In PT, these two morphemes belong to quite different procedures.

from the schedule in (2), because, unlike nominal morphemes such as possessive *'s* or plural *-s*, they are not generalisable, and their lexical selection is not obligatory but paradigmatic, that is, they are selected from a lexical list including other possible determiners such as article *the* or demonstrative *this* contributing to mark definiteness in N. Their lack of predictability and generalisability suggests that they do not qualify for inclusion in the schedule of morphological development. Nevertheless it is worth presenting the data in this (lexical) area of development given their parallel semantic role in the NP.

(5) *Possessive 's, and possessive determiners and pronouns in Kumi's corpus*

STRUCTURE	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14
POSSESSIVE (GENITIVE) 'S														
's		7			1>1		3			3	2		3>1	5>1
POSSESSIVE DETERMINERS														
MY		9	2	9		7	1	5	5	4	1	1	6	2
YOUR		5	1			1	1	1	7		8	4	8	11
HIS				1		1	1		8	13	1	3	6	21
HER				1				7	3	3	5	3	2	7
ITS														1
OUR			1						3	1				
POSSESSIVE PRONOUNS														
MINE				1			1		2				1	2
YOURS							1				1			
HIS												1		

As (5) demonstrates, Kumi's possessive determiners are numerous, consistent, and emerge at t2, that is, in parallel with possessive *'s*. In (6a) Kumi shows that, simultaneously with possessive *'s*, she is able to differentiate possessive determiners (*my big sister*) from SUBJ and other pronominals (*I go... me and my big sister... we play*). The utterance in (6b) shows that Kumi is well aware of the functional similarity between lexical possessive determiners and possessive *'s*, which she uses when the lexical item is missing at t5. Then at t9, when she has the lexicon, she progresses to the specific person determiner, as in (6c).

- (6) a. t2 I go . me and my big sister go to the Terry's house and we play
 b. t5 she hold she's frog
 c. t9 he call her name

The sequence in which possessive determiners emerge in Kumi's data shows that *my*, *your*, *his*, and *her* are the most numerous by far, followed by *our* and *its* far behind (5 and 1 occurrences respectively), whereas the 3rd person plural form

(*their*) does not occur in the corpus. Possessive pronouns, on the other hand, also occur, but they start much later than possessive determiners (at t4), are far fewer (10 occurrences in total vs 190 for possessive determiners), and comprise a smaller range of forms: mostly *mine*, a couple of occurrences of *yours*, and a single token of *his*.

Curiously enough, as mentioned above, possessive *'s* does not always appear in PT schedules and is rare in the published data. On the other hand, we have seen that it emerges early and is confidently deployed by Kumi, as her seven occurrences at t2 in (3a) and (5) testify. One may be tempted to say that Kumi's early acquisition and abundant use of possessive *'s* looks like L1 transfer from her native Japanese, whose possessive phrase has a similar frame (e.g., head last, genitive case marker on possessor), whereas learners with other L1s – such as Italian and Spanish (head first and no case markers on N), and Johnston's informants – seem to avoid it. With Pienemann, Di Biase, Kawaguchi & Håkansson (2005) we assume that transfer is developmentally moderated. So it is quite plausible that Japanese L1 learners are facilitated in learning that particular English structure early, first because it requires an early processing procedure (i.e., category procedure) and second because their L1 has a similar structure.⁵ This particular morpheme then could provide an empirical testing ground for the Developmentally Moderated Transfer Hypothesis on learner data from a variety of L1s with and without genitive case marking.

The next marker to emerge at the categorial stage is plural *-s* for Ns. For English this means that learners must be able to differentiate first whether the referent is considered as one entity or more (*cat* vs *cats*), and secondly whether the entity referred to is semantically countable or not (*apples* vs *milk*). Learners further need to learn that English requires this *-s* ending form also for referring to generic countable entities (*I like apples*), but not with generic uncountables (*I love coffee*) – not as simple as the learner might hope for, after all. This is especially so for Kumi, whose L1 does not mark plural morphologically. Kumi's data shows these difficulties. The plural *-s* marker emerges at t4, one month later than *-ing* and possessive *'s*, and a pattern of nonsuppliance persists up to t14⁶ as the table in (3a) above shows. In terms of structures belonging to the category procedure stage, this indicates the existence of a soft barrier. We note further that t4 is also the time when the phrasal procedure begins to be activated.

5 Japanese is left-branching whereas English is generally right-branching, but in this particular structure (i.e., possessive case marking on N) English behaves in a left-branching way (cf. Radford 2004: 169), like Japanese.

6 Which Ns fail to be marked with plural *-s*, e.g., generic versus non-generic, is worthy of further investigation (cf. Charters, Dao & Jansen 2011).

The next emerging structure, the so-called regular past *-ed* marker, provides a stronger case of a soft barrier, in so far as it shows up rather late in Kumi's corpus (at t9, about 18 months after arrival in Australia). This is rather late compared to 'irregular' past forms such as *came* and *went* used at t4, which are learned item by item. Once learned, the *-ed* marker is generalised to Vs that do not require it, as in (7a), and to newly learned Vs, whether regular or irregular as in (7b). Then at t14 (the 100th week of exposure) the child seems to have the rule under control. This developmental path is common for all learners of English L1, L2, and bilingual L1.

- (7) a. t9 he felled [/féld/] down
b. t13 the snail creeped up the jar

We defer further discussion of this apparent delay of the emergence of regular past *-ed* to where we discuss higher morphological procedures, except for saying that, in some ways, her delay reflects the difficulties Jia & Fuse (2007) found with this morpheme in Chinese L1 child learners in New York followed longitudinally for 5 years. For a treatment of its emergence in bilingual children with a variety of other languages, we refer to Nicoladis, Song & Marentette (2012).

Phrasal procedure stage. The next higher morphological procedure is the phrasal procedure, which assembles in a grammatically compatible way the various components of this syntactic unit. Its activation allows for intraphrasal agreement or, in any case, for exchange of grammatical information occurring within the phrase. The structures in the phrasal procedure stage seem to emerge in two steps. In Kumi's data, and perhaps this is true for ESL more generally, the first phrasal structure to emerge is the VP composed by the AUX *be* and a V marked with *-ing*. The activation of the *-ing* marker (recall that this is the earliest emerging morpheme) is gradually associated with forms of AUX *be*, with the earliest attempts occurring already in the earliest sessions, as (8a-d) show. These AUX *be* forms are soon produced in contrast with other AUXs (*can* and *can't*) emerging at about the same time and in a similar developmental curve, as the remaining examples (8e-k) show and the two rows of phrasal VP in the (3a) table indicate. So there is now the need for an exchange of information at the VP node between the AUX and the lexical V in order to produce either *V-ing* or a bare infinitive according to whether AUX is a *be* form or *can/can't* respectively.

- (8) a. t1 boy is hold the rock
b. t2 boy is go walk
c. t2 bird egg is er . falling down t1

- d. t3 he is go up
- e. t3 everyone can read this
- f. t4 this is fly (showing with her arms an aeroplane flying)
- g. t4 girl is eating the long chocolate
- h. t4 you can't open the door
- i. t5 dog is look the a hole
- j. t5 dog is look the bee nest
- k. t7 then they can thinking to play

Then Kumi's series of modal AUXs expands to include *will* and *couldn't* at t9-t10, as illustrated in (9).

- (9) a. t9 they will melt when they are hot
- b. t10 he couldn't see it

In sum, Kumi rarely misses unification in the VP, that is, in total 7 occurrences out of 119 – exhaustively shown in (8a-b, d, f, i-k) – and only three times after emergence.

The kind of agreement (government) which characterises the VP is grammatically stipulated and conceptually opaque. The ESL learner must learn to select AUX according to a range of aspectual, temporal or modal motivations, and unify these features with the relevant ones in the lexical V (*is going* vs *can go* vs *has gone*, etc.). Once this differentiation emerges it is practically categorical. Notice that, in Kumi's data, the structure *have* + V-*en* does emerge with some high frequency irregular Vs and contracted AUX forms, as shown in (10), but it does not occur for regular Vs. A proper PT treatment with an empirical investigation of the development of this set of VP structures in English requires more space than can be afforded here. Interested readers are referred to Yamaguchi & Kawaguchi (2014).

- (10) a. t11 they had put A and B and C to know which is which
- b. t13 no she's got a book
- c. t14 we've done this

The next step in the phrasal procedure stage is characterised by unification of features at the NP node when a plural context is created by numerical or nonnumerical quantifiers which must be unified with plural referents (*two cats*; *many dogs*). For example, the NP in (11), as well as the head N *cats*, contains a DET, a numeral quantifier and an adjective. The NP procedure here ensures that the head N is appropriately retrieved in its plural form *cats* compatibly with the numeral quantifier *two*, which is semantically plural, because English grammar requires such exchange of information; that is, the PL(ural) value of the head N's NUM(ber) feature must be compatible with any modifying element within the same phrase

which also has a NUM feature in its lexical entry.⁷ This example shows that, unlike Chinese or Japanese (which do not encode number morphologically), English encodes plural morphologically in countable Ns such as *cat*, but unlike Italian or Spanish it does not encode it in the adjective. Hence *black* does not display a number feature.

(11) a. *the two black cats*

b. lexical entries

the	DET	SPEC	"the"
two	QUANT	PRED	"two"
	NUM	PL	
black	A	PRED	"black"
cats	N	PRED	"cats"
	NUM	PL	

According to the PT acquisition criterion the child acquires plural *-s* without quantifiers (category procedure) at t4, then with numeric quantifiers at t7, and later with other quantifiers at t9 (both phrasal procedure). Agreement between N and its DET is also clearly intraphrasal. In English, this affects only the demonstratives (*this* vs *these*, and *that* vs *those*) and some quantifiers in plural forms (e.g., *lots of*). Kumi acquires phrasal agreement within NP at t7, and then gradually consolidates it with less frequent lexical items such as *most* at t9, as shown in (12), until she masters it completely at t14.

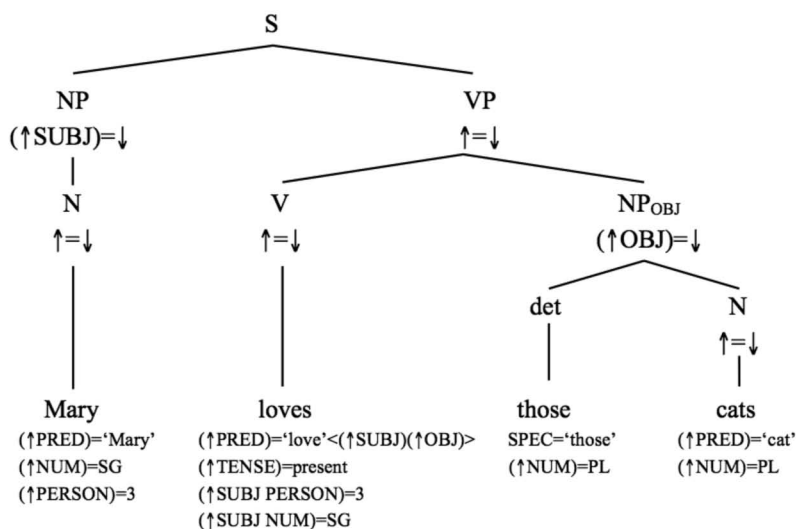
(12) t9 *most kids like it*

Recall that VP agreement is well in place at t4, that is, 8 months earlier than NP agreement, and soon becomes categorical. NP plural on the other hand is not only later, but it also takes longer for accuracy to be established, like the plural *-s* at category procedure level, with inaccuracies persisting to the end of the data collection. This is a clear case of a soft barrier, where various features seem to interact with number itself, such as definiteness, specificity, mass versus count, generic versus non-generic, as well as the possible influence of the L1. Since our database represents a single case study the patterns emerging here deserve further investigation over a greater number of learners, from different L1s and also covering a range of age groups.

⁷ This type of unification – motivated by a semantic feature in the NP environment, as mentioned in ch. 1, § 4.1, this volume – is clearly different from the other two types of unification, which are ‘concord’ and ‘government’; however, like the other two, it must be constructed online.

S-procedure stage. The activation of the next procedure, the S-procedure, allows for long distance, sentential (interphrasal) agreement between SUBJ and the lexical V. At this stage, learners of English can produce the verbal *-s* morpheme for the 3rd person singular of the present tense, once they manage to merge the SUBJ feature information in the NP_{SUBJ} (PERS = 3rd; NUM = SG) with the relevant V feature specifications (TENSE = PRES; SUBJ PERS = 3rd; SUBJ NUM = SG), as the diagram and lexical entries in (13) illustrate.

(13) *C-structure with lexical entries for Mary loves those cats*



At this stage, English morphology interacts with syntax, in the sense that, because SUBJ must agree with V (unlike OBJ), in order to produce this agreement, learners must identify one NP (the first) as SUBJ. This signals that functional assignment is in place. Kumi acquires the verbal 3rd person singular *-s* morpheme at t9, as shown in (14).

(14) t9 all the time he wakes up

Interestingly, this is the same time as she learns the past *-ed* morpheme. It may well be, then, that until the morphological marker *-s* on V (for tense, as well as person) emerges, there is little call for a systematic difference between another tense marker, that is past *-ed* and nonpast verbal forms. This in fact does not mean that Kumi never marks the past before she can activate the S-procedure, because she does use several irregular forms, such as *came* and *went*, already mentioned above, or *saw*

and *put*. After t9 more of the expected overgeneralisations of *-ed* (Pinker & Prince 1988) are produced but by the end of the data collection period Kumi realises it at 100%. This late emergence at t9 of the categorial-level tense marker *-ed*, well after the emergence of phrasal unification (i.e., *be V-ing*) at t4, does not constitute a threat to the general PT schedule for morphology because another lexical-level morpheme (*-ing*) has already emerged. In other words, not all of the possible markers belonging to a particular stage need to emerge in order to establish whether a particular PT stage has been achieved (cf. § 5, ch. 1, this volume). This is an often misunderstood point about PT's architecture.

S-BAR procedure stage. Further up the developmental sequence, we include here the interclausal level with subordination phenomena which in English affect morphological form, as the examples in the appropriate cell show (e.g., *I suggest he eat less* and *it's time you left*). These highly optional constructions showing an obligatory bare or a marked form of V in the subordinate clause belong to the 'educated' register of English. They are rather rare even in native speaker production and quite difficult to elicit in learners. Not only are they not produced by the child Kumi, but to our knowledge they are not yet documented in learner data. Needless to say not all subordinate clauses belong here. Neither, of course, is this the full story, in more ways than one, and several important issues are not mentioned here. For example, the activation of S-BAR procedure also includes tag questions and indirect questions, requiring interclausal agreement between V in the main clause and the subordinate clause. In any case, questions – as we argue – require a further pragmatic motivation, a focus, which is marked (morpho)syntactically (and prosodically) in the sentence, and as such are best treated separately.

In sum, Kumi's data shown in (3a) supports the developmental hypotheses in (2) with a calculated scalability of 100% if we consider only the three procedures, and 55% if we consider all the eight morphological structures shown in (3b). The latter result is well below other PT results in this volume and elsewhere, and indeed it is not valid, because the minimum coefficient of scalability is conventionally set at 60% (Hatch & Lazaraton 1991). This failure is almost exclusively due to the soft barrier imposed by the later emergence of *-ed* as a lexical morpheme marking past tense. Unlike the other V morphemes (*-ing* and 3rd per SG *-s*), which are almost completely regular forms, past *-ed* competes with a strong cohort of so called 'irregular' past V-forms, some of which are also far more frequent than any *-ed* form. This delays the emergence of the 'regular' form presumably on account of higher costs incurred by learners in the course of lexical selection of the appropriate form because they need to 'suppress' the competing and more frequent irregular patterns. This means that general morphological progression needs to take into account not only unification and syntactic distance but also lexical selection costs, which con-

firms the need for positing something such as our proposed soft barriers. Alternatively placing *-ed* at a higher stage, as Krashen (1977) does on account of relative timing, is not an option for PT, which does not relay on timing to place structures into stages but on whether or not the morpheme requires unification and at what syntactic level. Having said that, if we exclude the *-ed* structure from the calculation of scalability, we obtain a coefficient of 75% for the other seven morphemes over the three stages. This means that it is highly probable that they will be acquired in that order.

In concluding this part for morphological development, we may say that Kumi's data, as well as confirming both the universal PT schedules and the specific English ones, contributes towards refining and strengthening the general theoretical framework by providing additional evidence for internal hierarchies within each procedure. Thus, for the reasons discussed above in our schedule in (2), differently from Pienemann's schedules (e.g., 2005, 2011a, b), we place possessive *'s* below plural *-s* within the category procedure stage, and we 'populate' the VP slot within the phrasal procedure stage with a range of AUX V structures, and place it ahead of the NP agreement.

3. Syntactic development

In line with the presentation of the universal syntactic schedules afforded by the Prominence Hypothesis and the Lexical Mapping Hypothesis in chapter 1, § 4.2, this volume, we deal here with the learner's progress in English, thus testing the hypotheses on Kumi's longitudinal corpus first from canonical to noncanonical word order (§ 3.1), and then from default to nondefault mapping (§ 3.2).

3.1. *The Prominence Hypothesis*

Just as we keep the morphological hierarchy separate from the syntactic hierarchy on account of their different motivation, so, in dealing with the path the learner follows in acquiring the English syntax, we keep declarative sentences separate from questions. As Bettoni and Di Biase mention in chapter 1, § 4.2, the reason is that questions are pragmatically marked sentences, in so far as their universal defining characteristic is the focusing of the requested information, which must then be marked linguistically. This means that we will have three schedules based on the Prominence Hypothesis: one for declarative sentences, dealing with the DF TOP, and two for interrogative sentences (Y/N questions, and constituent questions), dealing with the DF FOC. Let us start with declaratives.

Declaratives. The syntactic hierarchy for English declarative sentences, based on the Prominence Hypothesis, is hypothesised in (15), and the distribution of the relevant syntactic structures over Kumi’s data is presented in (16).

(15) *Developmental stages hypothesised for L2 English syntax based on the Prominence Hypothesis: declaratives (after Pienemann, Di Biase & Kawaguchi 2005)*

STAGE	STRUCTURE	EXAMPLE
NONCANONICAL WORD ORDER	OBJ SUBJ V	<i>ice cream she likes</i>
XP _{TOP} CANONICAL WORD ORDER	TOP _{ADJ} SVO	<i>tomorrow they go home in Australia people eat pies</i>
CANONICAL WORD ORDER	SVO	<i>Mary jumped he working John eat rice</i>
LEMMA ACCESS	single words formulas	<i>station, here my name is Pim</i>

(16) *Kumi’s syntactic development based on the Prominence Hypothesis: declaratives*

STRUCTURE	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14
OBJ SUBJ V										-2			-1	1
TOP _{ADJ} SVO		6	2	6	7	3	2	3	4	6	5	3	5	16
SVO	19	27	37	42	23	33	21	24	42	20	29	16	25	22

As we saw for the development of morphology, the very first stage, with practically no grammatical marking, requires only the activation of the lemma access procedure. Again, as with the morphology data in (3a), the single word stage in syntax is taken as acquired by Kumi before t1, because in the first recorded session she already produces 19 canonical word order sentences.

Then, once learners can activate the category procedure and begin to differentiate verbal from nominal elements the stage is set for them to start organising their utterances using the language-specific canonical word order found in the input from their target language. Kroeger (2004: 141) offers criteria for deciding what is the unmarked or most basic word order in a language, among which he lists: highest frequency, widest distribution, neutral in terms of mood, polarity or voice (statements rather than questions or commands), positive (not negated) and active (not passive). This means that learners of English map concepts and mean-

ings by default onto a fixed NV sequence, that is, participant (mapped on N) before action or state (mapped on V). In grammatical terms this is a Specifier/Topic followed by Head/Comment order. When there are more than one participants in the event, the second one – being typically inanimate or in any case lower in topicality (cf. Dalrymple & Nikolaeva 2011) or conceptual hierarchy – is placed after the head, like any other predicative or circumstantial material such as time/place. This establishes an NVN sequence which most often will correspond to sequences of actor-action followed by patient/theme and/or location and/or time. Provided they are affirmative, minimally presuppositional, and pragmatically neutral, these sentences are syntactically target-like, as shown in (17). Here we report the very beginning of a continuous stretch of Kumi's production when she is telling a frog story from the book *Frog where are you?* at t1, that is, 4 weeks after her arrival in Australia. Notice that, unlike Pienemann (2011b: tab. 4.7, p. 63), we do not consider *then* in a narrative as an adverbial ADJ, but as a sentence connector (conjunction) similar to *and*. Notice also that actions or states may be expressed by lexical items other than V, such as a movement preposition (*down*) in (17i), and that the NV sequence is produced most of the time, and can be assumed to function as the syntactic core of the expressed meanings.

- | | |
|---|--|
| (17) a. and boy sleep er | [N V: Actor Action] |
| b. and frog go out. of the hor house | [N V PP: Actor Action Place] |
| c. and boy wake . | [N V: Actor Action] |
| d. no frog in bottle | [no N PP: (Neg) Theme Place] |
| e. then boy look the under shoe | [N V PP: Actor-Action Place] |
| f. and dog er in | [N pause False start] |
| g. the dog look er in the bottle | [N V PP: Actor Action Place] |
| h. and boy open the a window
and and shout | [N V N and V: Agent Action Theme & Action] |
| i. then dog down er window | [N P/V N: Actor Action (movement)] |

The next step forward in the development of syntax accounted for by the Prominence Hypothesis is achieved in English when the learner's lexicon and categorial differentiation are strengthened and expanded, and TOP is placed in first position. This TOP constituent is not typically a core element of the sentence – that is, it can be a time or place circumstantial ADJ in declarative sentences, to which prominence is assigned by the speaker. This structure comes about early and is target-like⁸ in a highly configurational language such as English, where canonical word order is undisrupted in declarative sentences that are pragmatically rea-

8 This same kind of operation would produce ungrammatical sentences in German, which requires V in second position (cf. ch. 9, this volume).

sonably neutral – we say ‘reasonably’ here, because even the topicalisation of ADJ is pragmatically marked, albeit minimally compared to the topicalisation of a core constituent other than SUBJ, which is also a DF and TOP by default when no other GF is marked as TOP (cf. the notion of Stage Topic in Erteschik-Shir 2007, Shibatani’s (1994: 277–8) ‘stylistic focus’, and the discussion on the nature of SUBJ in Falk 2006). Kumi reaches this intermediate stage at t2, as (16) and (18) show.

- (18) a. t2 today we need a hat
b. t2 in kindly we have a rain drop

At this stage, which we call the ‘TOP_{ADJ} canonical order’ stage, the canonical string remains unaffected by topicalisation. However, speakers may still need to attribute prominence to grammatically more central constituents, typically OBJ. At first, when OBJ is topicalised, the Prominence Hypothesis proposes that learners keep the canonical string complete by duplicating the topicalised element in its canonical position with a pronominal element. Kumi does exactly this twice at t10 and once more at t13 in a kind of structure-preserving operation, as (19) shows. In fact the syntax of all three examples – two in (18) at t2, and one in (19) at t13 – may be represented with the single phrase structure rule in (20), where S = canonical word order.

- (19) t13 tomato thing I like it

- (20) S' → (XP) S
(↑TOP) = ↓ ↑ = ↓

At the last stage, native-like noncanonically aligned sentences are hypothesised, allowing for highly marked constructions such as those rarely used in English (cf. Bresnan’s example reproduced in ch. 1, (19b), this volume) – being highly configurational, English prefers other than syntactic means for assigning prominence to a particular element in a declarative, e.g., prosodic means involving stress and intonational contours (cf. Selkirk 1984; Levelt 1989: 303–305), a domain which remains unexplored within PT. This last syntactic stage emerges in Kumi only at t14 with a single example shown in (21). So this seems to be a large step forward, achieved late.

- (21) t14 this you can click. on lots of buttons

Interrogatives. As remarked before, questions are not pragmatically neutral constructions, and universally require the focusing of the requested – new and unpredictable – information (Kroeger 2004: 138–139). With Y/N questions, the focus

broadly falls on the whole sentence by default (wide focus), as in (22), whereas with constituent (or content, or *wh*-) questions it falls on a specific constituent (narrow focus), as in (23).

(22) a. did you have pizza for lunch?

b. yes

(23) a. what did you have for lunch?

b. a pizza

Specific constituents may of course be brought into focus also with Y/N questions, as the felicity of the answer in (24b) attests.⁹ In English, this sort of focusing may be achieved by prosodic means (Lambrecht 1994; Van Valin 2005), as shown in (24a), but we do not deal with this narrow focus here because, as we said, prosody is beyond the proposed scope of this chapter (and PT), as well as being out of reach for the early L2 learner, who would have to become highly familiar with default L2 prosodic patterns in the first place in order to use them in this flexible way.

(24) a. did Mary buy a CAR?

b. she bought a motorbike

Across languages, different means are used to tell interrogatives apart from declaratives, such as special intonation patterns, a particle or clitic, a modal affix on V or a change in word order (Kroeger 2005: 203–205). In English, the wide scope of Y/N questions has grammatical reflexes such as word order restrictions and other lexical properties of AUXs. These ordering and selectional restrictions are represented in the lexical entry of AUXs, and learners, however gradually, will need to acquire them in order to produce marked order questions, involving AUX-SUBJ order among other things. Also, Vs are capable of being assigned information structure (*i*-structure) features [\pm prominent] and [\pm new] (cf. ch. 1, 4.2.1, and (36), this volume). On the other hand, the narrow scope of constituent questions also has the same grammatical reflexes, and additionally it requires the *wh*-element bearing the pragmatic focus in initial position. In the following pages, with Kim & Sells (2008: 191, ch. 10), we adopt the feature question – QUE(stion) – for marking any of these exponents characterising questions, as well as Choi's (1999, 2001) *i*-structure features, without going into formalisation details. We will deal with polar questions first because they are more numerous in our informant's data (138 out of 180), and then with constituent questions.

⁹ This possibility of narrowing the scope of Y/N questions skews them towards content questions (cf. Givón 2001: 230–34).

Polar questions. Our proposed schedule for the development of Y/N questions is hypothesised in (25), and then tested on Kumi’s data in (26).

(25) *Developmental stages hypothesised for L2 English syntax based on the Prominence Hypothesis: Y/N questions*

STAGE	STRUCTURE	EXAMPLE
NONCANONICAL WORD ORDER	AUX _{QUE} SUBJ V (O)	<i>does he like pizza?</i> <i>have you tried pizza?</i>
	MOD _{QUE} SUBJ V (O)	<i>can Ann swim?</i>
	have _{QUE} SUBJ OBJ	<i>have you a boyfriend?</i>
	COP _{QUE} SUBJ predicate	<i>is Joan happy?</i> <i>are you there?</i>
QUE CANONICAL WORD ORDER	QUE [SVO]	<i>do they have cat?</i> <i>is your man have a red hat?</i> <i>is Mary is happy?</i>
CANONICAL WORD ORDER	[QUE ^p SVO]	<i>dog eating the doughnut?</i> <i>you like pizza?</i> <i>you are there?</i> <i>Tom is happy?</i>
LEMMA ACCESS	[QUE ^p single words] [QUE ^p formulas]	<i>Jim happy?</i> <i>coffee? going?</i>
QUE ^p = the feature is exclusively prosodic		

(26) *Kumi’s syntactic development based on the Prominence Hypothesis: Yes/No questions (re-analysis of Yamaguchi 2010)*

STRUCTURE	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14	
COP SUBJ PREDICATE						1							1	3	
AUX SUBJ V(O)						2				1				7	3
DO SVO			17	7	7	6	16	9	9					9	1
SVO (LEXICAL V)						1	3	1	2						
SVO (COPULAR V)			7				1								
SINGLE CONSTITUENT	1	5			1			5	3	1	2				6

In single-constituent Y/N questions the referent, known to both interlocutors, requires confirmation or further information. This requesting modality, formalised as QUE^P in (25), is marked by prosody exclusively (with rising intonation). Some of Kumi's examples are shown in (27).

- (27) a. t1 this?
 b. t2 light green?
 c. t2 blue coat?
 d. t4 long or short?
 e. t7 yes or no?
 f. t7 one man?

Polar questions at the next hypothesised stage (i.e., the canonical word order stage) are also expressed by prosodic means only, but now the learner of L2 English is able to form a full SVO string. So, in the very first two stages, the contrast between the QUE^P modality and the declarative modality is entrusted to prosodic means exclusively. And this is why, up to now, the modality marker QUE^P in (25) is enclosed in square brackets to indicate that it does not yet have a lexical exponent. Only later, at the QUE canonical word order stage, is a small set of lexical material (e.g., *do, do you, do you have, do you got*) preposed to the canonical string and added to the prosody to jointly express the question modality, and later still, at the non-canonical order stage, these lexical means are integrated, as AUX, within the sentential syntax. This lexically expressed QUE is not enclosed in square brackets and is, informationally, [+prominent]. We will now show these developments in some more detail as we look at Kumi's data.

In Kumi's data, structures belonging both to the [QUE^Pcanonical order] and QUE [canonical order] stages appear at the same time, quite early at t2. This does not contradict PT because there is a conflation, not a skipping of the stage. At the canonical word order stage, interestingly, the order NVX with copula appears earlier than with lexical V, as illustrated in (28). Quite likely this is aided by the fact that the copula is far more frequent in the input than any lexical V.

- (28) a. t2 your light is off?
 b. t2 bicycle is a red?
 c. t4 dog eating the doughnut?

What is even more interesting is the way in which, at the QUE canonical order stage, Kumi initially discovers that QUE may have lexical exponents expressed with *do you have* and *do you got*. Despite their occasional apparent grammaticality, the *do* element here is not yet integrated as a proper AUX within the questioned sentence. Significantly, at this time, Kumi's AUX system in the declara-

tive grammar is not fully developed either. Nor, of course, could the whole string *do you have* or *do you got* ever be an XP – as the general progress of the Prominence Hypothesis in chapter 1, § 4.2.1, would have it. This is indeed the very feature that sets polar questions apart from other types of sentences, and consequently the very reason why in our English Y/N question schedule in (25), the XP in prominent position is replaced by QUE in prominent position also. Kumi's examples in (29) support this proposition of a *do* QUE which at this stage is a lexical step towards developing a fuller AUX system. The 37 QUE elements produced by the child from t2 to t5 are all *do you have...*, some of which (e.g., 29a-c) look grammatical, whereas others (e.g., 29d-f) are not. Here Kumi is attempting to create 'syntactic amalgams' (Lambrecht 1988), main or subordinate, which are as yet beyond her grammar. A very similar progression towards the formation of complex clauses is observed also in English L1 acquisition (Diessel 2004).

- (29) a. t2 do you have a man on the picture?
 b. t2 do you have a flower on the leaf picture?
 c. t2 do you have a trash can in the picture?
 d. t2 do you have a man coat is a blue?
 e. t2 do you have a man wear the sun glasses?
 f. t2 do you have a flower is pink?

Later on, at t6, *do you have* enters in competition with *do you got*, as in (30a-c), which however is soon abandoned altogether (from t7) in favour again of *do you have*. Later still (at t9), a further development occurs when Kumi seems to understand the need to signal polar questions with a QUE AUX other than *do* by replacing it with an inflected *be* form, as in (30d).

- (30) a. t6 do you got a newspaper?
 b. t6 do you got a man wearing the red hat with yellow ribbon?
 c. t6 do you got a bin in there is a rubbish in bin?
 d. t9 is your man have a red hat?

Already at t6, however, a more important development seems to be happening when Kumi not only experiments with an alternative (ungrammatical) QUE (*do you got*) form, but also starts producing two Y/N structures belonging to the next marked alignment stage, which displays an integrated AUX, as shown in (31).

- (31) a. t6 can you see?
 b. t6 is bicycle falling down?

We can thus assume that from now onwards the ungrammatical use of the *do you have* QUE element followed by other V-like material declines, and that instead an integrated fronted *do* AUX becomes productive whenever there is no COP or other AUX element available in the sentence to mark the polar question.

After their initial production at t6, Kumi's Y/N questions belonging to the last noncanonical word order stage remain rare for some time (until t13); that is, she produces one COP example at t7 and one MOD AUX example at t9, as shown in (32).

- (32) a. t7 are you here?
b. t9 can I have the that walkman?

Finally, at t13 and t14, Kumi is able not only to identify *do* AUX fully as governing lexical V and carrying 3rd person singular SUBJ information, as in (33b), but also to include a new, and less common, AUX such as *could* in (33c).

- (33) a. t13 do you have a boy listening to music?
b. t14 does your man have green coat?
c. t14 could you ask me a question?

In sum, for polar questions, as the development of COP and AUX continues, it is possible to trace the learner's progress from a more lexical question marker followed by canonical order to a more legitimate AUX properly integrated in a sentence with marked word order. Torregrossa & Bettoni (2013) report similar results for Y/N questions in English EFL data by adult Italian learners.

Constituent questions. For a typological perspective on constituent question formation within the nonderivational framework of LFG, we refer to Mycock (2007). Constituent questions are realised typologically by two distinct means: in-situ languages, such as Japanese or Chinese, signal focal prominence of question phrases by prosodic means (and a specific lexical set), whereas fronting languages, such as English and Italian, signal it by syntactic means (and their specific lexical sets) – although this does not mean that prosody does not play a role (e.g., the *wh*-word is usually stressed in English), but simply that syntax plays a crucial role. In native speakers' English, however, in-situ questions such as those in (34) are perfectly acceptable given appropriate pragmatic context and intonational contour, although they are rare and highly marked, expressing for instance amazement, surprise or disbelief. In fact, although typologically languages are clearly either in-situ or fronting, neither of these means of forming questions is exclusive of the other in a single given language.

- (34) a. [Tom ate a fly] Tom ate what?
 b. [Anne is going to the South Pole] she is going where?

We will now lay out the developmental hypotheses for English constituent questions in (35), and then present Kumi's results in (36). Note that Kumi's L1, Japanese, is typologically an 'in-situ' language, and she is learning English, a 'fronting' language.

(35) *Developmental stages hypothesised for L2 English syntax based on the Prominence Hypothesis: constituent questions*

STAGE	STRUCTURE	EXAMPLE
XP _{FOC} NONCANONICAL WORD ORDER	WH _{QUE} AUX SUBJ V (O)	<i>what has Tom eaten?</i> <i>where did Joan go?</i> <i>when are you going?</i>
	WH _{QUE} MOD SUBJ V (O)	<i>what can Mary do?</i>
	WH _{QUE} COP SUBJ	<i>where are they?</i> <i>what is this?</i>
XP _{FOC} CANONICAL WORD ORDER	WH _{QUE} SVO	<i>what he eat?</i> <i>where Joan is?</i>
CANONICAL WORD ORDER	WH _{QUE} in-situ	<i>Joan eat what?</i>
LEMMA ACCESS	single words formulas	<i>what? where?</i> <i>how much is it?</i>

(36) *Kumi's syntactic development based on the Prominence Hypothesis: constituent questions (re-analysis of Yamaguchi 2010)*

STRUCTURE	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14
WH _{QUE} MOD SUBJ V (O)														2
WH _{QUE} AUX SUBJ V (O)									1		2			1
WH _{QUE} COP SUBJ X								1	3		1	8	6	4
WH _{QUE} SVO								1	1		1			
SINGLE WORDS; FORMULAS	1		1	7	3		4	1	1					2

Our informant Kumi produces 180 questions in total as we mentioned, of which content questions are less than one third (52). This imbalance may be due to the lack of specific tasks designed to elicit this sort of data. Also the range

of wh-words produced is restricted. For instance, among GFs, SUBJ is not questioned, and neither is any OBL.

At the lemma access stage Kumi produces one instance of *what* at t1 and a peak of this wh-word at t4 (3 occurrences), when it alternates with the formulaic single word (*I beg your*) *pardon?* (also 3 occurrences). From t7 onwards, Kumi uses the five question phrases listed in (37).

- (37) a. t7 how many flowers?
 b. t7 which one?
 c. t7 what the noise?
 d. t14 which grandma?
 e. t14 what flowers?

It may be argued that these are neither single words nor formulas, and thus perhaps requiring an intermediate stage before the WH_{QUE} canonical order stage. However, we do not propose this because Kumi's data is too limited, and leave them at the initial stage because they are without a V. The remaining 8 questions at this stage are clearly formulas because they are all exactly the same (*where are you?*) and coincide with the title of the book *Frog where are you?* which was well known to the child and used several times over the two years of data elicitation.

In our hypotheses for L2 English in (35), we place in-situ wh-questions at the next canonical word order stage. Also Pienemann (2007) places them within the hypothesis space of early learners. Yet it is interesting to note that, at an early stage, they are attested neither in the literature, nor in Kumi's corpus. That is, learners from a variety of language backgrounds appear to latch on quickly to the specific structural means prevalent in the target language, regardless of whether their own L1 may use a different mechanism. A case in point is Kumi, whose L1 Japanese prefers in-situ question. Yet, like other learners of English, she appears not to use them, and soon learns the English characteristic of fronted focal prominence, as her results will show. Similarly, children in bilingual L1 acquisition where English is the dominant language outside the home appear not to display in-situ constituent questions in English (Qi 2012; Qi & Di Biase 2005 for Chinese-English; Itani-Adams 2007, 2011 for Japanese-English). However, in other contact situations, such as that of Cantonese-English bilingual L1 acquisition in Hong Kong, meticulously described by Yip & Matthews (2007: 93-111), in-situ questions do show up strongly and unequivocally. This may point towards a significant role of environmental influences in language development (cf. the discussion for L2 Italian data in ch. 8, this volume).

At the next stage, our hypothesis is that having learned to disentangle the DFs TOP and FOC from SUBJ, learners of English can now place the focal questioned constituent, represented by the wh-word, in first position followed, steadily, by canonical word order. Kumi has three such questions, all *why* questions as shown

in (38). Here the XP is not an argument but an ADJ. Thus this constituent question structure with FOC_{ADJ} parallels the TOP_{ADJ} canonical order structure predicted at this same stage for declarative sentences by the Prominence Hypothesis – cf. (15) and (18).

- (38) a. t8 why you come and shout in my house?
 b. t9 why you fall down?
 c. t11 why I put the crab in the jar?

As the sentences in (38) show, the outcome with *why* (ADJ) questions is not target-like in English – that is, it would be if the questioned constituent were itself SUBJ, but we have already remarked that there is no *who* question in Kumi's data. SUBJ also has a special status in content questions in other languages besides English, according to Falk (2001). For a study on adult L2 English which includes *who* questions, we refer to Kawaguchi (in press). Her cross-sectional data supports Falk's claim in so far as *who* questions appear to be highly problematic for all learners except for the most advanced ones.

The developmental path for content questions concludes at the marked alignment stage, that is, when, after establishing the question word or phrase in FOC position, the learner is able to use functional assignment and place SUBJ in the appropriate post-COP/MOD/AUX position. Kumi starts off producing constituent questions belonging to this stage with the *wh*-word followed by COP at t8. In (39) we show Kumi's four examples occurring at t8 and t9. Recall that polar questions with COP were already established at t6. So the addition of the *wh*-FOC does not alter the post-AUX position of SUBJ.

- (39) a. t8 what's the purse?
 b. t9 where is she?
 c. t9 where are you? [twice]

As the reader may have noticed, Kumi's two questions in (39c), coinciding with the title of her book, are exactly the same as those we have placed among the formulas when she produced them earlier on. However, there is evidence that she can now construct them online. The evidence for this is provided by her other two examples: (39a), where COP not only occurs before SUBJ but is also contracted; and (39b), where the same 3rd person COP form is not contracted and the SUBJ pronoun is correctly in 3rd person in contrast with the *are* in (39c).

Constituent questions belonging to the marked alignment stage with the *wh*-word before AUX emerge in Kumi's corpus at t9. Notice that this AUX is actually a form of *do* (40a-c), which has a strong pattern of preceding the SUBJ in polar questions already at t2. In the very last session of data collection the child concludes

her developmental path and produces two examples of the more sophisticated wh-word with modals (40d-e) which were also established earlier (at t6) as preceding SUBJ in polar questions.

- (40) a. t9 what do you means?
 b. t11 what did she do?
 c. t11 what did they do?
 d. t14 what would he do?
 e. t14 what should they do?

In concluding this section on the development of syntax according to the Prominence Hypothesis, we should again warn the reader – as we have already done with respect to the development of English morphology – that this is not the full story for questions. For example, although usually included in presentations of PT, indirect questions and tag questions are not dealt with here. Our reason for leaving them out is that they both involve two clauses, the former through subordination, and the latter through long distance unification of verbal features across coordination; furthermore the latter also involve negation. Indeed, negation and its complex interaction with questions is not treated at all here. These constructions require specific investigations dealing with subordination and negation.

3.2. *The Lexical Mapping Hypothesis*

As we have seen in chapter 1, § 4.2.2, this hypothesis deals with the mapping of a-structure to f-structure, accounting for the development from default to nondefault structures, as motivated mainly by lexical selection, which may in turn be based on discourse-pragmatic motivations. The universal schedule shown there in (43) is language-specifically applied to L2 English with illustrative examples provided here in (41).

This staged development for L2 English has been convincingly tested in several empirical cross-sectional studies, including one by Wang (2010). Another, also by Wang (2006, 2011), uses a large body of data gathered by means of patient-cued pragmatic contexts, and demonstrates that only the most advanced Mandarin-speaking learners of English are able to produce sentences using the pragmatically appropriate nondefault syntactic mapping. A third study by Keatinge & Keßler (2009) investigates passive constructions in classrooms. A fourth by Kawaguchi (2011) involving Japanese L2 investigates the acquisition of nondefault mapping (e.g., passive and causative as well as unaccusative and psychological Vs) which may not be observed frequently in naturalistic, longitudinal studies. We further test this staged development for L2 English on Kumi's longitudinal data here in (42) including the intermediate stage of default mapping with additional argument.

(41) *Developmental stages hypothesised for L2 English syntax based on the Lexical Mapping Hypothesis*

STAGE	CONSTRUCTIONS	EXAMPLES
NONDEFAULT MAPPING	unaccusatives, passives, causatives, exceptional verb constructions, etc.	<i>the vase broke</i> <i>Bob was beaten by Ted</i> <i>she made him cry</i> <i>we received a letter</i>
DEFAULT MAPPING AND ADDITIONAL ARGUMENTS	agent/experiencer mapped on SUBJ, patient/theme mapped on OBJ, and other members of the a-structure hierarchy, such as goals and locatives, mapped on OBL	<i>Mary put the butter in fridge</i> <i>she gave Tim a new bike</i> <i>Ann went to Rome by train</i>
DEFAULT MAPPING	agent/experiencer mapped on SUBJ patient/theme mapped on OBJ	<i>John sleeping</i> <i>John fry egg</i>
LEMMA ACCESS	single words formulas	<i>station, here</i> <i>my name is Pim</i>

(42) *Kumi's syntactic development based on the Lexical Mapping Hypothesis*

STRUCTURE	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14
unaccusatives														1
passives											(1)		1	
agent/experiencer on SUBJ, goal on O ₁ , and theme on O ₂						1			1	1				
agent/exper. on SUBJ, patient/theme on OBJ, and other arguments on OBL						12	4	2	1	1				
agent/experiencer on SUBJ	6	12	6	4	9	5	8	8	12	3	10	6	5	4
agent/experiencer on SUBJ, patient/theme on OBJ	13	15	31	38	14	15	9	14	28	15	19	10	20	18

The number in brackets indicates that the structure is ambiguous: it could be a passive or an adjectival predicate.

In dealing with the development of declarative sentences in § 3.1, we have already shown the early steps taken by Kumi in mapping arguments in her new L2. In the examples in (17) it is clear that initially the actor/agent-like roles are the ones most commonly mapped in TOP (or SUBJ) position with theme/patient mapped in OBJ position.

Not before t6 (after about four months' exposure) does Kumi reach the intermediate stage of 'default mapping with an additional argument' where the added argument is marked as such, that is, distinguishing it from SUBJ and OBJ, which are marked by position only. At t6, this additional element is marked both as OBL by a preposition and as OBJ₂ by position, as in (43a) and (43b) respectively.

- (43) a. t6 owl came from a tree hole
 b. t6 we give the kangaroo a kangaroo food

Unfortunately Kumi's longitudinal study stops when nondefault mapping is just emerging in her interlanguage. This last stage is only represented in her data by the two examples shown in (44), one produced with a passive structure at t13, and the other with an unaccusative V at t14.

- (44) a. t13 the cat was chased by the mother bird
 b. t14 is your bike not broken and **fall failed over**?

Nevertheless we can clearly see that her data fully supports our schedule hypothesised in (41).

4. Conclusion

In this chapter we have offered a refreshed and in some areas more complete and supported map of L2 English within a PT perspective. For morphology, we clarify the status of possessive $\rightarrow s$ within the categorial stage and propose to exclude possessive determiners and pronouns, identifying them as lexical rather than morphological structures. In addition we propose to place the unification of AUXs with lexical Vs firmly within the phrasal stage. We have also exemplified and, at least in part, explained the scouting and trailing observed in intrastage sequencing by proposing the concept of soft barriers to account for those instances where feature unification and syntactic distance may be insufficient to account for previously unexplained progressions within stages.

For syntax, we coherently incorporated and further explored the intuitions broached in Pienemann, Di Biase & Kawaguchi's 2005 PT extension, leading to our proposed Prominence Hypothesis, which embraces questions as well as declaratives. We believe that the longitudinal data examined here, despite some gaps, supports our current vision. Given the limited data belonging to a single child learner it would be important to validate and complete our proposals with further data with a broader set of learners in a wider age range. It will also be crucial to pursue an appropriate formalisation incorporating i-structure in a wider Optimality-theoretic LFG (OT-LFG) launched by Bresnan (e.g., 2000) (cf. Asudeh & Toivonen 2010: 453).

Needless to say, there are still gaps in this story. Several important issues are not discussed here. For example, the activation of S-BAR procedure yields a further stage in English, which includes, for instance, tag questions and indirect questions, requiring interclausal agreement between V in the main clause and V in the

subordinate clause. In any case, questions – as we argued – require a further pragmatic motivation, a focus, which is marked (morpho)syntactically (and/or prosodically) in the sentence, and as such are best treated separately. Another area not discussed here includes the variation implicit in the Hypothesis Space (Pienemann 1998: ch. 6) and also the relationship between accuracy and development. As for the Lexical Mapping Hypothesis we are barely scratching the surface here, but we believe we are on the right track and would like to encourage other scholars to take this interesting area further.

3

The development of Italian as a second language

Bruno Di Biase* and Camilla Bettoni**

*University of Western Sydney, **University of Verona

1. Introduction

Italian is a nonconfigurational, null-SUBJ headmarking language characterised by a rich morphology and a flexible syntax which is highly sensitive to pragmatic and discourse choices.¹ From the point of view of the effect of pragmatics on syntactic structure, Van Valin (2005: 77) locates Italian among languages with ‘flexible syntax and rigid focus’. English, on the other hand, is among languages with ‘rigid syntax and flexible focus’, which makes the contrast between the two languages intriguing. These typological characteristics are of interest to PT in two fundamental ways. First, with regards to the notion of transfer of grammatical information within and between phrases of a sentence (cf. ch. 1, § 4.1, this volume), Di Biase & Kawaguchi (2002) show that, despite the basic contrast with English, Italian interlanguage data fully validates the universal hypotheses about the development of morphological structures and their interaction with syntax as hypothesised in Pienemann (1998), who had not looked at any Romance languages. Secondly, and perhaps even more importantly, with regards to the LFG architecture of correspondences among its three parallel levels of linguistic representation, the need to account for the nonconfigurationality of Italian syntax contributed substantially to the formulation of PT’s hypotheses about the development of syntactic structures at the interface with discourse-pragmatics (cf. ch. 1, § 4.2, this volume). As a matter of fact, Di Biase & Kawaguchi (2002) pioneered the use of the newly formalised LFG DFs in PT, thus foreshadowing the extension later developed in Pienemann, Di Biase & Kawaguchi (2005). In what follows, unlike any previous treatments of Italian processability, we revisit and expand the morphosyntactic framework for Italian L2 development and propose a theoretically motivated way

1 For a brief description of Italian grammar written in English, cf. Maiden (1995), and Vincent (2011). For more comprehensive treatment in Italian, cf. Renzi, Salvi & Cardinaletti (2001), Salvi & Vanelli (2004), and Schwarze (2009) – the latter within an LFG framework.

forward for dealing with the so-called intrastage phenomena (cf. § 2). We also offer a fairly comprehensive discussion of the interface between syntax and discourse-pragmatics, with empirical support (cf. § 3).

2. Morphological development of L2 Italian

PT-derived hypotheses for the acquisition of Italian must deal with its rich morphology instantiating all-pervasive and obligatory agreement patterns.

In terms of morphological typology, Italian is located higher than English on the index of fusion continuum. This is the index which measures the extent to which morphemes are segmentable, with agglutination at one end, where segmentation is straightforward, and fusion at the other end, where there is no segmentability (Comrie 1989: 46). English morphemes are more easily segmentable than Italian morphemes, which in most cases fuse a number of grammatical features in a single exponent. Segmentation of inflectional morphemes is often more problematic in Italian even than in other Romance languages such as French, Spanish or Portuguese, all of which, for instance, have adopted suffixation of *-s* to mark plural in nominal inflection. Italian on the other hand has a system of vowel alternation (Vincent 2011), which makes nominal number and gender hard to factor out, and more opaque for learners.

The other important characteristic of Italian morphology is that it is stem-based, like Russian and Hebrew, rather than word-based, like English or German. This is significant from a processing point of view, because – for the vast majority of nouns and adjectives, and for all verbs – Italian stems do not amount to full legal words, and must always bear some inflectional ending. The function of these inflectional endings is to express grammatical categories such as number, gender, mood and tense (Maiden 1995: 92). For example, the lexical item in (1a) cannot be realised in its bare stem (1b), but it must have one of the four inflectional vowel endings typical of Italian nominals, as in (1c). The inflectional endings in (1c) mark the gender contrast (masculine vs. feminine: e.g., *ragazzo*, ‘boy’ vs *ragazza*, ‘girl’; *ragazzi*, ‘boys’ vs *ragazze*, ‘girls’) and the number contrast (singular vs. plural, e.g., *ragazzo*, ‘boy’ vs *ragazzi*, ‘boys’; *ragazza*, ‘girl’ vs *ragazze*, ‘girls’) in nominals. Learners appear to acquire the phonological part of the process very early, namely, the fact that Italian words typically display a vocalic ending. But then it takes them much longer to account for the grammatical information loaded in the vocalic variation they hear in the input at the end of words.

- (1) a. {*ragazzo*} [boy]
 b. */ragats-/
 c. /-o -a -i -e/

Apart from the irregularities found in any system, nominal group marking in Italian is made more complex than the paradigm presented in (1) by the existence of several phonologically-based noun classes. In addition, from a semantic point of view, Ns with features +human and/or +animate do not always match their 'natural' and grammatical genders. All other nouns are assigned by the grammar to one or the other gender in an arbitrary way, often following phonologically based criteria: e.g., Ns ending in the unmarked singular citation form *-o* tend to be assigned to masculine gender (*libro*, 'book'), and those ending in *-a* to feminine gender (*casa*, 'house'). Yet Ns ending in *-e* are masculine (*pane*, 'bread'; *leone*, 'lion') or feminine (*neve*, 'snow'; *tigre*, 'tiger') in an arbitrary way.

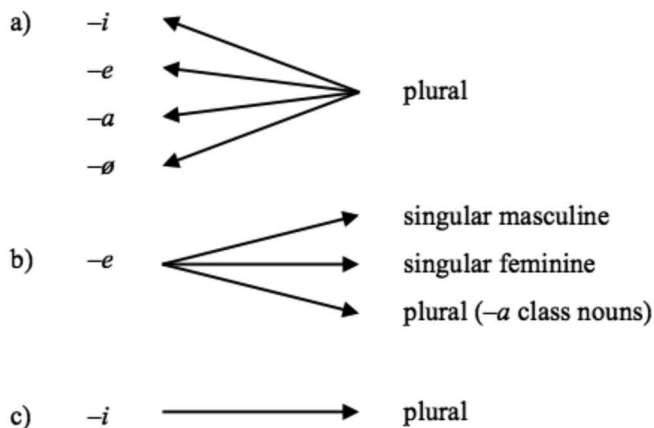
Also nominal modifiers, such as determiners, demonstratives and adjectives, must express the same gender and number values as their head Ns. Nominal modifiers also fall into classes: those with the same four endings seen in (1c) for Ns (*rosso* - *rossa* - *rossi* - *rosse*, 'red'), and those which neutralise the gender distinction by having *-e* ending for singular, and *-i* ending for plural (*verde* - *verdi*, 'green'), irrespective of whether their head is masculine or feminine. An anchor point in this extreme variation is offered by the stability of *-i* as appearing with plural referents consistently. No wonder, as we shall see, that learners latch on to this morpheme and use it as a kind of prototypical, or default, plural marker. This phenomenon of identifying a certain form as bearing some particular function by default seems to be pervasive across all levels of morphological as well as syntactic and pragmatic marking.

The task faced by the learner in sorting out Italian nominal inflection is complex enough. Yet it is rivalled by that imposed by verbal morphology. We will not deal with the latter here, except to mention briefly that Italian Vs fall into three classes, each with a characteristic thematic vowel distinguishing three conjugations (*-a-*; *-e-*; *-i-*); and that a typical V has 47 or so finite forms, marking tense, aspect and mood, as well as person and number.

The complexity of the Italian inflection system offers a good example of the way in which the primary PT notion of information exchange within and across constituents needs to be complemented by other principles in order to explain the acquisitional process. Among these, there is the form-to-function relationship (Pienemann 1998: § 4.3). That is, the actual learning of the morphological form of the affix in relation to its function is a different task from that of managing information distribution in the affixation process, where diacritic features have to be exchanged within different grammatical structures. The figure in (2) illustrates how Italian Ns mark the *plural* value of their number feature through a complex set of form-function relations. The many-to-one relationship, where several morphemes mark one and the same feature is exemplified in (2a). Then there is the one-to-many relationship, as shown in (2b), where a particular morpheme marks more than one function. On the other hand, the most consistent

form-to-function relationship is shown in (2c), where the final *-i* vowel at the end of Italian nominal forms only marks the number feature [plural]. Similar form-function mapping problems may also be expected with the acquisition of Italian verbal paradigms, where the vowel ending of one form (e.g., *mangia* “(s/he) eats/is eating”) carries information regarding several features at once, such as subject person, subject number, tense, aspect and mood. A veritable labyrinth for the learner.

(2) *Form-function relations of Italian plural nominal markers*



In essence, the relationship between morphological forms and their functions exhibits different degrees of complexity. This adds another dimension to the learning task which is separate and different from the task on which PT is focused, namely the exchange of grammatical information and the use of diacritic features. So far PT has not made any predictions on how a fuller paradigm develops. However, on the one hand, the more regular and simpler one-to-one form-function relationships may help to bootstrap the more complex ones (cf. Andersen's 1984 “one-to-one principle in interlanguage construction”). On the other hand, teasing out of different factors allowing to progress from emergence² to full mastery of the whole system is one of the directions in which future research can go – and not only with regards to Italian (cf., e.g., the brief mention of Russian morphology in ch. 5).

2 For a thoughtful assessment of the emergence criterion in PT as applied to Italian, cf. Pallotti (2007).

2.1. The hypotheses

Let us now consider some of the main language-specific Italian L2 structural outcomes of the morphological processing procedures universally predicted by PT (cf. ch. 1, § 4.1, this volume). Our hypotheses are shown in (3), and then discussed with interlanguage examples taken from Di Biase's corpus analysed in Di Biase & Kawaguchi (2002) and Di Biase (2007), and briefly described in § 2.2.

(3) *Developmental stages hypothesised for L2 Italian morphology (extended from Di Biase & Kawaguchi 2002)*

STAGE	MORPHOLOGICAL OUTCOME	STRUCTURE	EXAMPLE
S-BAR PROCED.	INTERCLAUSAL MORPHOLOGY	subjunctive marking in subordination	<i>immagino siano partiti</i> [I imagine they have left]
SENTENCE PROCED.	INTERPHRASAL MORPHOLOGY	NP _{TOPi} Cl _{OBJi} AUX V-to _i NP _{SUBJ}	<i>i fichi li ho comprati io</i> [the figs, I have bought them]
		NP _{TOPi} Cl _{OBJi} V NP _{SUBJ}	<i>i fichi li compro io</i> [the figs, I buy them]
		NP _{SUBJ} AUX V-to	<i>i bimbi sono partiti</i> [the kids have left]
		NP _{SUBJ} COP predicative adjective	<i>i bimbi sono buoni</i> [the kids are good]
PHRASAL PROCED.	VP MORPHOLOGY	AUX V-to	<i>sono usciti</i> [(they) have left]
		COP predicative adjective	<i>sono buoni</i> [(they) are good]
	NP MORPHOLOGY	N adjective	<i>bambini buoni</i> [good children]
CATEGORY PROCED.	LEXICAL MORPHOLOGY	person marking on V	<i>mangiare vs mangio</i> [eat vs I eat]
		past marking on V	<i>mangia vs mangiato</i> [eat vs eaten]
		plural marking on N	<i>bambino vs bambini</i> [child vs children]
LEMMA ACCESS	INVARIANT FORMS	single words	<i>no lavoro</i> [no work]
		formulas	<i>mi chiamo Karim</i> [my name is Karim]

After leaving behind the single-words and formulaic (lemma access) stage, learners begin to incorporate language-specific procedures at the next (lexical-morphology)

stage, when categorial marking for Italian Ns is achieved through the use of article-like forms, such as *la* or *il* as generalised markers (cf. also, in a non-PT framework, Chini & Ferraris 2003: § 3.4.1), as *la acqua* ‘the water’ in (4), and *l’oca* ‘the goose’ in (5) further below.

- (4) Researcher: sotto nell’acqua? [under the water?]
 Trish: no no la acqua [no no water]

In PT the category procedure is characterised by the lack of information exchange with any other element in the phrase or clause. Our interpretation then would not consider the combination of noun and article as evidence for phrasal procedure morphology, independently of whether or not the specific combination turns out to be target-like (cf. also Di Biase 1998). In fact, bare Ns are not often produced at early stages of Italian L2 development, and are in any case highly restricted also in native Italian, aside from personal names. Furthermore, the Italian article is syntactically incapable of appearing by itself and is prosodically not independent of the content word that follows, and groups under its stress field. This, in turn, would favour a sort of formulaic (exemplar) learning of article-N combinations. For all these reasons, such combinations are hypothesised as belonging to the lexical level, in the sense that article forms are considered as part and parcel of categorial marking of Ns rather than as agreeing determiners³. A clearly different situation obtains with other determiners or N modifiers such as demonstratives, which may appear (syntactically) independently or at the end of phrase boundaries, and have their own independent stress pattern and prosody. When such N modifiers must agree with their N head, a phrasal procedure needs to be called upon.

At this category procedure stage in Italian, the plural *-i* diacritic, being the most consistent marker of plurality in the language among several other markers (cf. (2) above), turns out to be the first to emerge in conjunction with plural referents, as in (5).

- (5) Lois: l’oca ritornato *caroti
 the goose returned carrots
 [the goose has returned the carrots]

Like Ns, Vs also show categorial marking in Italian, such as the infinitive *-re* ending used in various contexts, as in (6). At this stage, perfective past tense can also

3 Naturally, some defaults may be learned quickly (*la casa*), others (*il problema*) may take a long time (cf., the discussion on soft barriers below).

be marked by the *-to* past-participle ending,⁴ although not yet in analytical constructions with their AUX, as in (5) above. Some person marking on V is also attested at this stage. See, for example in (6), the formal contrast between *capire* and *capisco* that may mark first person with the characteristic null SUBJ. This contrast, however limited, is sufficient to show that person marking appears at a much earlier stage in Italian compared to English – a fact that can be explained by the null-SUBJ nature of Italian (cf. our discussion of this point below).

- (6) Lois: non... capire... non capisco
 not understand-INF not understand-1.SG
 [(I) don't... understand... (I) don't understand]

The morphological processes that characterise the next, phrasal, stage in Italian interlanguage include nominal and verbal agreements. Within the NP, learners start producing the agreement of determiners (other than articles as we have seen), and/or adjectives in attributive function, with the gender/number of the head N, as in (7). Within the VP, unification of number value (singular or plural) produces two types of agreement: one between the copula and a predicative adjective or a nominal, as *sono cugini* in (8); the other between the person marker in the (*essere*) AUX⁵ and the ending in the lexical V, as in (9).

- (7) Anne: non ho tanti amici maschili [maschi]
 not have-1.SG many-PL.MASC friends-PL.MASC male-PL.MASC
 [I don't have many male friends]

- (8) Amy: sono cugini della mia mamma
 are-3.PL cousins-PL.MASC of my mother
 [they are my mother's cousins]

- (9) Toni: ah sì, sono andati alla camera di letto
 oh yes are-3.PL gone-PL.MASC to the bedroom
 [oh yes, they went to the bedroom]

Let us now move on with the learners to the S-procedure stage. Notice that in English, on account of obligatory SUBJ, person variation in the V-form is placed

⁴ The *-to* marker also marks perfective aspect but it is difficult to segment out tense and aspect marking.

⁵ Analytical tenses require different AUXs (*essere* 'be', *avere* 'have') that carry the finiteness features (Schwarze 2009: 150). However only the *essere* AUX is relevant for phrasal agreement because it requires GEN/NUM agreement between SUBJ and lexical V.

high in the processability hierarchy (cf. ch. 2, § 2). Italian, on the other hand, being a null-SUBJ language, maps the person-number (singular or plural speaker, addressee or third person) directly on V form without a necessary co-reference to a nominal or pronominal SUBJ (cf. Di Biase & Kawaguchi 2002 for an LFG formalisation). Indeed SUBJ may not be expressed at all, or be generated after V (cf. § 3.1 below). Results from psycholinguistic experiments (e.g., Vigliocco, Butterworth & Semenza 1995; Vigliocco, Butterworth & Garrett 1996) support the hypothesis that SUBJ-V agreement in null-SUBJ languages is generated via an independent retrieval of the features of V and those of SUBJ. If this is the case, then – for Italian and other null-SUBJ languages – achievement of interphrasal morphology may be more clearly expressed by structures other than SUBJ-V agreement, so as to allow for the fact that at least some of the different person-number forms of V are acquired, as we have just seen, at an earlier stage.

In Italian too, of course, interphrasal morphology requires the S-procedure, that is, the procedure for unifying different categories of constituents at sentence- or clause-level. This means that, for the emergence of structures belonging to this stage, the learner must recognise the grammatical relations (e.g., SUBJ, OBJ) expressed by the various constituents of the clause, as well as identify the category of each constituent, and more generally the relationship between predicates and their arguments, including predicates of an adjectival or nominal nature, as we have seen in (8)-(9). So what are the candidate structures for Italian at the sentence agreement stage? One structure that can be built on those in (8)-(9) is the unification of SUBJ features (gender and number) with nonverbal predicates, as in (10).

- (10) Anne: *i genitori di mia mamma sono australiani*
 the parents-PL.MASC of my mother are-3.PL Australian-PL.MASC
 [my mother's relatives are Australian]

Other good candidates include agreements in verbal analytic constructions (with AUXs) that are likely to be unified online, provided they require nondefault (i.e., not singular and masculine) unification. By this we mean not the unification of the person feature of SUBJ, which is carried by the AUX, but of the values for its number and gender features, which must be unified with the lexical V, as in (11).

- (11) Amy: *noi siamo andate da Napoli a Palermo*
 we-1.PL are-1.PL gone-PL.FEM from Napoli to Palermo
 [we went from Napoli to Palermo]

Here, the number value of the lexical V form (*andate*, 'gone') is unified with that of the pronominal SUBJ (*noi*, 'we'): plural in both cases. On the other

hand, the feminine gender value also marked on the lexical V (*andate*) is not marked in the pronominal SUBJ (*noi*), which could indifferently refer to males, females or mixed referents. So, we may ask, where does the feminine gender information of the lexical V-form come from? The answer to this question must be that – because both features (gender and number) are required by V – the gender value is retrieved by the V lemma directly from the conceptual structure. Pronominal SUBJ, on the other hand, requires only the number value. It is these kinds of feature distributions and unification patterns that lead Di Biase (2007: § 1.2) to support the ‘independent retrieval’ assumption of Vigliocco and her co-workers, who carried out numerous experiments concerning SUBJ-V agreement in a range of typologically different European languages (e.g., Vigliocco, Butterworth & Garrett 1996; Vigliocco, Hartsuiker, Jarema & Kolk 1996; Vigliocco & Franck 1999, 2001). This line of research suggests that, at least in null-SUBJ languages with SUBJ-V agreement, both SUBJ and V retrieve features from the conceptual structure independently, and then merge them at the S-node. Hence, the V-form is essentially a lexical-stage form, stored in the mental lexicon of the speaker, whose features must match or be compatible with its agreeing counterpart.

Still at the S-procedure stage, and again referring to (3) above, another structure hypothesised for L2 Italian is the TOP-V agreement occurring in clauses that topicalise OBJ by (dis)placing it to the left of V from its canonical postverbal position (cf. § 3.2 on syntactic development). In such cases OBJ is an accusative clitic pronoun co-referential with the TOP placed at the beginning of the clause. These (i.e., the NP TOP and the OBJ clitic) must agree in number and gender values, as in (12). Furthermore, if the V is in an analytic construction with an AUX, its past participle form will bear the same number and gender values as the clitic, as in (13). This structure then requires that learners recognise a full nominal TOP as nonSUBJ, and mark their discourse-pragmatic choice explicitly by the ACC clitic pronoun exhibiting NUM and GEN values in agreement with the NP TOP. Learners who can produce such long distance agreement must clearly be able to assign SUBJ and OBJ functions, and manipulate their agreement and position patterns. More about this complex structure will be said when presenting the development of Italian syntax in § 3.1, and in chapter 8.

- (12) Toni: i broccoli li compra il cane
 the broccoli-PL.MASC them-PL.MASC buys the dog
 [the broccoli, the dog buys it]

- (13) Amy: le patate le ha comprate la il cane
 the potatoes-PL.FEM them-PL.FEM has bought-PL.FEM the dog
 [the potatoes, the dog has bought them]

Finally, at the S-BAR procedure stage, subordination phenomena at the interclausal level can affect morphological form also in Italian, as they do in English (cf. § 2, ch. 2 in this volume). In (14) and in the appropriate cell in (3) above there are a couple of examples.

- (14) speriamo venga domani
 hope-INDICATIVE.1.PL come-SUBJUNCTIVE.3.SG tomorrow
 [we hope she comes tomorrow]

These constructions, which show a marked form of V in the subordinate clause, belong to an educated register of standard Italian; they are rather rare in native speakers' everyday production, and difficult to elicit in learners. As already noticed with regard to L2 English in chapter 2, learner data has yet to support the late emergence of this last stage for L2 Italian. Indeed, the very complex area of subordination is still a research gap in PT, and requires further theoretical elaboration and focused empirical investigation.

2.2. *Hard barriers and soft barriers in morphological development*

Evidence in support of the learner's journey outlined in § 2.1 is provided by two studies: Di Biase & Kawaguchi's (2002) cross-sectional study of six English-speaking Australian university students learning L2 Italian, two each attending beginner, intermediate and advanced courses; and Bettoni & Di Biase (2005) longitudinal study of a seven-year old Romanian girl recorded for eight times after her arrival in an Italian school. In (15) we extend the analysis presented in Di Biase & Kawaguchi (2002) by adding further phrasal and interphrasal structures, and focus on the morphosyntax involving the most frequent, consistent and reliable form-to-function marking and agreement phenomena. The oral production data set from the six learners (and one native control interacting with the researcher, not shown in the table) was elicited over two sessions totalling between 35 and 60 minutes for each learner. The first session included free conversation, a picture description and a story telling task. A shorter second session focused on a communicative task devised by Di Biase to elicit structures with OBJ clitics, which he called the 'animal dinner task'. This task shows the learners two cards at a time, one with the picture of animals, and the other with that of food items. When the two cards appear on the computer screen, learners are encouraged first to plan who (the agent) is buying what (the theme) for a forthcoming dinner in the (simple) present tense, and then to retell the events by saying who has brought what in the (analytic) past tense. Learners have to do this starting from the card on the left. Since this card randomly shows the animals and the food items, active SVO

(15a) *Cross-sectional study of morphological development in L2 Italian (extended and updated from Di Biase & Kawaguchi 2002)*

STAGE	STRUCTURE	TRISH	LOIS	CARRIE	ANNE	TONI	AMY
S-PROCED.	NP _{TOPi} Cl _{OBJi} AUX V-to _i NP _{SUBJ}					-6	+5-2
	NP _{TOPi} Cl _{OBJi} V NP _{SUBJ}		-19	-11	-1	+11-1	+12-1
	NP _{SUBJ} AUX V-to			+1	+6-3	+2-2	+10-6
	NP _{SUBJ} COP predicative adjective	-1		+1-1	+6	+3-2>1	+16-3
VP PROCED.	AUX V-to				+3-2	+3-2	+4-1
	COP predicative adjective		-3	+3-1	+10-1	+6-5>1	+3-1
NP PROCED.	N adjective	-1	+1-5	+3-1>1	+7-1	+4-3>1	+6-2
CATEGORY	past -to marking on V	+1-2	+10-9	+16-3	+26	+112-11	+54-3
PROCED.	plural -i marking on N	+1-14>5	+14-5>1	+12-2>2	+21	+17-2>1	+10
+ = supplied in obligatory context (correct) > = oversupplied in a singular context (error) - = not supplied in obligatory context (error) empty cell = no context							

(15b) *Scalability matrix of Italian morphological data shown in (15a)*

STAGE	STRUCTURE	TRISH	LOIS	CARRIE	ANNE	TONI	AMY
S-PROCED.	NP _{TOPi} Cl _{OBJi} AUX V-to _i NP _{SUBJ}					⊖	+
	NP _{TOPi} Cl _{OBJi} V NP _{SUBJ}		-	-	⊖	+	+
	NP _{SUBJ} AUX V-to			-	+	+	+
	NP _{SUBJ} COP predicative adjective	-		-	+	+	+
VP PROCED.	AUX V-to				+	+	+
	COP predicative adjective		-	+	+	+	+
NP PROCED.	N adjective	-	-	+	+	+	+
CATEGORY	past -to marking on V	-	+	+	+	+	+
PROCED.	plural -i marking on N	-	+	+	+	+	+

sentences (e.g., *l'oca compra le carote*, 'the goose buys the carrots') should alternate with topicalised ones (e.g., *le carote le ha portate l'oca*, 'the carrots, [them] has brought the goose') – although it is of course impossible to rule out the use of passives, a more formal choice in Italian (e.g., *le carote sono comprate dall'oca*, 'the carrots are bought by the goose'). The total data set thus gathered consists of about 30,000 words, of which half produced by the learners, with a mean length of 10.3 words per turn.

The initial lemma access stage is omitted from the results because it is safely reached by all learners. The rest of the counting is exhaustive over the database. However, default agreements (e.g., the citation, and most frequent, masculine/singular form) are excluded because otherwise the acquisition criteria would be clouded over by the bulk of the default itself, as we will explain where it becomes relevant. For the organisation of the table, see the criteria in chapter 1, § 5, this volume.

Our empirical results for all six learners in (15a-b), with nine implicationally arranged structures generating a matrix of 54 cells, support the hypotheses formulated in (3), yielding a coefficient of scalability (Hatch & Lazaraton 1991: 204-16) of 100% if we consider only the three procedures. This means that the principle of 'information exchange' and its operation across major linguistic units (categorical, phrasal and interphrasal) is robust. In chapter 1, § 5, we called 'hard barriers' these boundaries that the learner needs to negotiate from one stage to the next in order to be recognised as having achieved a particular developmental stage. Thus Trish has not yet crossed the hard barrier to the category procedure stage. Lois has, but has not crossed into phrasal procedure stage yet. Carrie successfully reaches the phrasal procedure stage but not completely. The other three learners can be said to be capable of handling S-procedure but, curiously, each of them seems to be at distinct points of the interphrasal stage. So, having said that the hypothesis is supported, if we consider all the structures severally, our data yields a coefficient of scalability of 89%, which is more than sufficient to reduce the possibility of chance in the sequence in which these structures emerge.

How can one interpret this difference between a scalability of 100% for structures collapsed within the stages and a lower scalability for them taken individually? Notice that a similar pattern obtains also for English, as shown in (3a), § 2, ch. 2, this volume. The answer is probably that, whereas there are categorical implicational relationships among the procedures described by PT, these relationships are not as strong between specific structures belonging to each stage. A refinement of the gross division by stages was foreshadowed by Mansouri & Håkansson (2007) under the guise of 'intrastage' sequencing in learner language. These scholars identify structures belonging within a single stage as being acquired according to a specific pattern. More specifically, three

different structures for marking definiteness in Arabic and two in Swedish, respectively, were found to emerge in a specific sequence. Mansouri & Håkansson (2007: 115) then propose an expanded 'Hypothesis Space' model for PT to account for "more complex linguistic and functional phenomena that have implications for language processing". Furthermore, from a methodological point of view, they indicate the desirability of identifying for any stage the 'optimal structure' for empirically testing PT-based predictions, because the structures emerging later within a stage may not be good candidates for testing theoretical predictions. We concur with Mansouri and Håkansson that it is desirable to find an 'optimal' structure to characterise the stage. A case in point is the *-i* plural marking for Italian, the most frequent and reliable out of all the possible markers for plurals. However, we propose that it is crucial to identify the source(s) of complexity, so that the sequencing itself also becomes theoretically predictable (and falsifiable).

There are two main sources of language-specific variation that constrain the learner's progress over these barriers within a stage, which we called 'soft barriers' (cf. ch. 1, § 5, ch. 2, § 2, for English, and ch. 4, § 3.2, for Japanese). These sources are the lexicon, namely, the bundling of lexical features and lexical mapping, on the one hand, and on the other hand, the discourse-pragmatic requirements expressed in the grammar, as originally identified in the PT extension (Pienemann, Di Biase & Kawaguchi 2005). Both lexical features and discourse requirements are hypothesised to absorb further processing cost according to the number of syntactic nodes requiring unification and, consequently, requiring greater attention by L2 speakers (cf. Segalowitz 2003; Taube-Shiffnornan & Segalowitz 2005). These two constraints can be seen at play in the results presented in (15), where the solid line represents the 'hard barrier' between one stage and the next, and the zigzag line represents the 'soft barrier' of a hypothesised step within the stage. In other words, while the environment of feature unification (i.e., the scope of information exchange) provides the basic constraint between one stage and the next, both the quality and quantity of the features requiring unification over two or more nodes and discourse requirements (which require greater attentional resources) will further constrain progress within a stage. Let us consider these two sources in turn.

With regard to the language-specificity requirement of particular lexical items, focusing on our results in (15), we notice that the accuracy rate for some morphological agreements at the higher stages is still poor with the more advanced learners. Why is this so? On the one hand, some of the structures identified as candidates for our hypotheses rely on analytical constructions based on nontransitive Vs selecting AUX *essere*. In fact, along with Maiden (1995: 150) and Schwarze (2009: 153), among others, we consider active transitive constructions which select AUX *avere* as default given their frequency,

consistency and high predictability.⁶ Because *avere* constructions do not require agreement with the lexical V, they are not very informative in terms of developmental stages in morphological agreement. On the other hand, there is a small group of Italian Vs with unaccusative meaning (e.g., *andare*, *partire*, *nascere*) that select *essere* AUX ('to be') in analytic forms (Maiden 1995: 150–153). In these cases, as in copular constructions and with passives and reflexives, the participle form of V must agree in gender and number with SUBJ. This type of concord is a laborious thing to learn and may motivate what we call here a soft barrier within the stage for the learner to overcome, causing the rather shaky accuracy we can see in the results of more advanced learners such as Anne, Toni and Amy. Their lower rates in fact occur when attempting structures with *essere* AUX (unaccusatives, passives, reflexives) in an otherwise strong agreement performance. The intermediate stage learner, Carrie, interestingly crosses over the hard boundary to achieve the interphrasal stage, with agreements between copula and predicative adjective, but she does not produce any contexts for the *essere* AUX and V-to agreement. We agree with a reviewer of this chapter that the evidence here is weak because no production does not necessarily imply no acquisition; however, all learners have performed exactly the same tasks designed to elicit this structure and only the three more advanced learners produced it successfully.

The effect of the second source of language-specific variation constraining the learner's progress within a stage can be exemplified at the top of the S-procedure stage. Because all of the three more advanced learners have crossed over the hard barrier between phrasal and interphrasal stages, the difference between them can be characterised by the soft barrier imposed by the appearance of topicalised OBJ. That is, Anne does not topicalise, although she attempts to use some passive constructions, as in (16), which, however, do not involve clitics.

- (16) Anne: mm la lattuga è comprata dal cane
 the lettuce-SG.FEM is bought-SG.FEM by the dog
 [the lettuce is bought by the dog]

Toni crosses the OBJ topicalisation boundary, as (15) shows, but does not manage the long distance TOP-V agreement over the next soft boundary, which Amy does, as (13) shows. It could be argued that the computation of the discourse-pragmatic information together with the morphosyntactic informa-

6 Some intransitives (e.g., unergatives such as *dormire*, *camminare*) also select *avere*, and do not require agreement. Like transitives, these Vs tend to have an agentive SUBJ.

tion may indeed require the S-procedure but that TOP lay outside of it. In other words, it could be argued that there should be a 'hard' boundary to separate more clearly the discourse-pragmatic information from the rest of the S-procedure because, even though the linguistic unit is still the clause, in processing terms the long distance agreement is computed over potentially discontinuous constituents. For the moment it may be best to let the 'linguistic unit' define the hard boundaries, and let the additional lexical or discourse information define the soft boundaries (some of these issues will be picked up again in § 3.1). In any case, the coefficient of scalability of the same results for the same learners shown in (15) improves dramatically from 89 to 100% when soft barriers are taken into account.

3. Syntactic development

Italian appears to assign a lesser role to syntax than to morphology in interpreting GFs (e.g., SUBJ can be null, but always marked in verbal morphology). As Bresnan (1998: 119) observes more generally, morphological forms will compete with and pre-empt phrases that carry no additional information. If the syntactic structure nodes do not bear additional GFs that distinguish them from the morphological structures, they must be omitted. This explains why the numerous Italian word order options are used more for mapping pragmatic and semantic information than for conveying grammatical information.

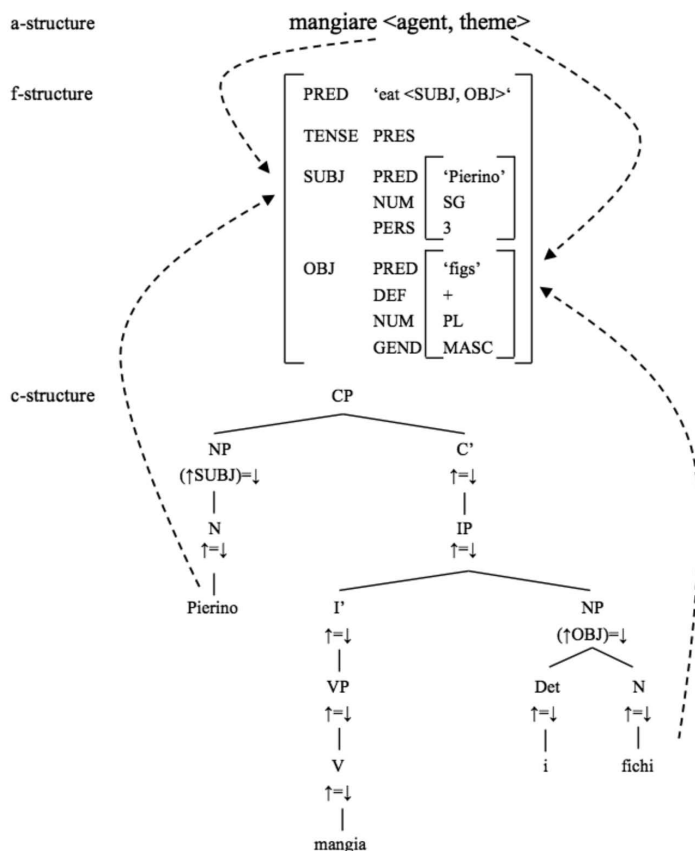
3.1. *The Prominence Hypothesis*

Let us then look at how the Prominence Hypothesis applies to the development of Italian syntax. We will present first the key features of Italian grammar on which the predicted developmental trajectory is based, and then the actual trajectory for L2 Italian declarative sentences, drawn from work by Di Biase and his collaborators (e.g., Di Biase 2005; Di Biase & Bettoni 2007; Bettoni, Di Biase & Ferraris 2008; Bettoni, Di Biase & Nuzzo 2009; Bettoni & Di Biase 2011). The development of Italian content questions is discussed and tested with empirical data by Bettoni & Ginelli in chapter 8, this volume.

Like English, Italian canonical word order is SVO, as shown in (17). This means that, in pragmatically neutral, simple, declarative sentences, SUBJ is the default TOP, and OBJ is the default FOC.

However, since Italian – unlike English – is a nonconfigurational language, its canonical word order can be freed up for mapping pragmatic and semantic information. In theory at least, all permutations of the core elements are possible, as shown in (18), where the propositional meaning of all the sentences is 'Pierino eats

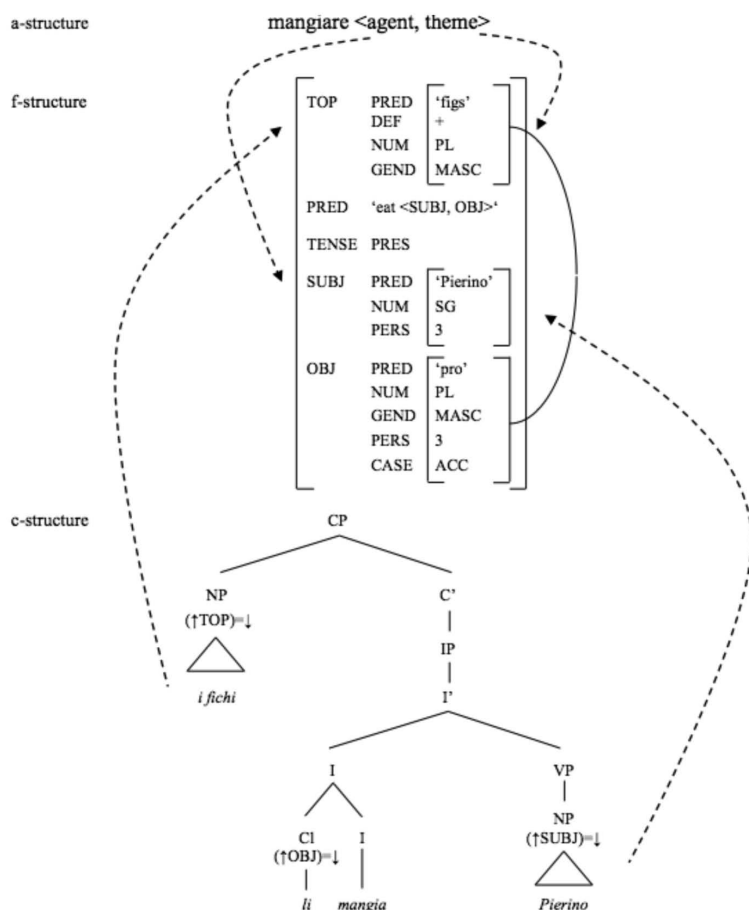
(17) *Canonical correspondences between a-structure, f-structure and c-structure for the sentence Pierino mangia i fichi* ('Pierino eats the figs')



the figs'. In practice, however, those in (18c-f) are highly dependent on marked prosody for interpretation and rarely used. OBJ topicalisation, as in (18b) and illustrated formally in (19), needs no special prosodic emphasis, because the "functional uncertainty" of the first NP (Bresnan 2001: § 4.8) is solved by a coreferential OBJ clitic marker on V (both *i fichi* and *li* are MASC-PL).

- (18) a. SVO Pierino mangia i fichi
 b. OVS i fichi li mangia Pierino
 c. SOV Pierino i fichi mangia
 d. OSV i fichi Pierino mangia
 e. VSO mangia Pierino i fichi
 f. VOS mangia i fichi Pierino

(19) *Noncanonical correspondences between f-structure and c-structure for the sentence i fichi li mangia Pierino ('the figs, Pierino eats them')*



If the inanimate nature of the first NP in this sentence will semantically rule out the possibility of the *figs* doing the eating, confusion could easily arise if animacy is shared by both the participants in the eventuality. This would happen in (20), for example, if neither SUBJ nor OBJ is unequivocally marked. In configurational languages such as English one can tell positionally which NP has which GF: in declaratives, the one before V is SUBJ, and the one after V, if any, bears some other GF. As we have just seen, Italian marks preverbal OBJ with a clitic agreeing with TOP: in (18b) masculine and plural (*li* and *i fichi*), and in (20) masculine and singular (*lo* and *il bambino*).

- (20) a. la mamma accarezza il bambino
 SUBJ V OBJ
 [mummy caresses the child]
- b. il bambino lo accarezza la mamma
 TOP.MASC.SG he-CL.ACC.MASC.SG V SUBJ
 [the child, mummy caresses him]

If, on the other hand, it were *the child* doing the caressing (as in 21), and for discourse-pragmatic reasons the speaker would wish to place this NP in postverbal focal position and place *mummy* in topical position (as in 21b), then the clitic signalling all this would agree with *mummy* and be feminine and singular.

- (21) a. il bambino accarezza la mamma
 SUBJ V OBJ
 [the child caresses mummy]
- b. la mamma la accarezza il bambino
 TOP.FEM.SG she-CL.ACC.FEM.SG V SUBJ
 [mummy, the boy caresses her]

Besides being nonconfigurational, Italian is also a null-SUBJ language. In spoken Italian, SUBJ is neither referential nor pronominal in almost 70% of sentences (Bates 1976), and is expressed exclusively by verbal morphology. This means that, when previously mentioned, implied or shared, SUBJ is left out, as in (22a). Should pronominal SUBJ be used, it would indicate emphasis or contrast, and therefore most often occupy the postverbal focal position in declarative sentences,⁷ as in (22b), rather than its canonical preverbal position.

- (22) a. hai sentito i ragazzi? hanno telefonato
 have-2.SG heard the boys? have-3.PL phoned
 [have you heard from the boys? they have phoned]
- b. hai chiamato tu i ragazzi? hanno telefonato loro
 have-2.SG called you the boys? have-3.PL phoned they
 [was it you who has called the boys? it was them who have called]

⁷ Because Italian allows for flexible word order, it is a non-plastic language, i.e., it does not allow variable placement of nuclear stress (Vallduvi & Engdahl 1996), in contrast to English, for example.

Even from this brief presentation of Italian word order rules, it is easy to see that learners will be able to acquire them all only gradually. Let us then illustrate their path hypothesised for syntactic development with regard to word order (cf. 23).

(23) *Developmental stages hypothesised for L2 Italian syntax based on the Prominence Hypothesis: declaratives (after Bettoni & Di Biase 2011)*

STAGE	STRUCTURE	EXAMPLE
NONCANONICAL WORD ORDER	TOP _i Cl _{OBJ,i} -V SUBJ	<i>i fichi li compro io</i> [the figs, (I) buy them]
XP _{TOP} CANONICAL WORD ORDER	TOP _{ADJ} SVO	<i>al lavoro non ho amici</i> [at work (I) have no friends]
CANONICAL WORD ORDER	SVO	<i>Pino mangia fichi</i> [Pino eats figs] <i>parlo italiano poco</i> [(I) speak Italian a little]
LEMMA ACCESS	single words formulas	<i>ciao</i> [hello] <i>mi chiamo Ugo</i> [my name is Ugo]

After the initial stage, when single concepts are mapped to single words or formulas, learners will at the next stage organise words according to the order most typically and most frequently recurring in the L2 input. This is SVO for Italian, including the possibility of VO, or V, as we have seen. Notice however that, at this stage, the canonical-order sentence remains underspecified regarding the grammatical functions of its core referents. That is, learners will analyse the preverbal NP, if present, semantically as agent or pragmatically as TOP rather than purely grammatically as SUBJ.⁸ Recall also our comment on the categorial stage in morphology in § 2 above: formal variation on V may begin to emerge in Italian interlanguage at this stage – that is, much earlier than in English interlanguage. So, whereas the English *-s* marker on third person singular emerges when the learner is able to unify the relevant features of SUBJ-V agreement at the interphrasal stage, a variety

⁸ As a matter of fact, given that SUBJ is claimed to be universal (cf. the Subject Condition of LFG, Falk 2006: 98, 170), learners assume that one element has SUBJ status. This would mean a mapping by default between SUBJ, TOP and agent. However, at this stage SUBJ remains underspecified, with some of its Italian-specific properties unrealised.

of morphological verbal markings begin to appear at this category stage, because – thanks to the null-SUBJ character of Italian – there is no SUBJ for the V to unify with. The sentences in (24)-(25) are typical of this stage.

- (24) mia famiglia. fratello sorella mangia dolce swedish
 my family brother sister-SUBJ eat-V sweet swedish-OBJ
 [my family, brother sister eat swedish sweet]

- (25) mangio sul lavoro
 eat-V at work-ADJ
 [I eat at work]

To move beyond the canonical word order stage, learners must assign grammatical functions to sentence constituents independently of their fixed position within the canonical string. PT's Prominence Hypothesis predicts that they will begin to do this by uncoupling TOP from SUBJ. The first step in this development happens when learners begin to contextualise (in time, space, etc.) the core sentence by adding an ADJ to canonical order. If for pragmatic reasons this addition is introduced as TOP occupying a position early in the sentence as in (26), there will now be two preverbal constituents. This will force learners to distinguish between TOP and SUBJ, change the relationship between the c-structure and f-structure, and require additional processing procedures. With the addition of ADJ as TOP, however, the first position is occupied by a nonargument constituent, so the SVO structure nudges intact to the right.

- (26) mezz'ora dopo mio padre prende il treno
 half an hour later-TOP_{ADJ} my father-SUBJ catches-V the train-OBJ
 [half an hour later my father catches the train]

Further in their development, learners will also be able to disrupt canonical word order. Once again this will happen gradually: first when only one core argument is in a noncanonical position, namely when SUBJ is postverbal, as in (27), and then when both are assigned to a noncanonical position, namely when SUBJ is postverbal and OBJ preverbal, as in (28).

- (27) compro io il pane
 buy-V I-SUBJ the bread-OBJ
 [I am going to /it's me buying the bread]

- (28) hindi lo sanno tutti
 Hindi-TOP_{MASC.SG} it-CLACC.MASC.SG know-V everybody-SUBJ
 [Hindi, everybody knows it]

Supporting evidence for this trajectory comes from several corpora, gathered in Australia and in Italy. We summarise here the results of the analysis of a cross-sectional data set of 15 learners working or studying in Verona or Milan (Bettoni & Di Biase 2011). They are all in their twenties (9 females, 6 males), with a wide range of competence levels in L2 Italian and L1 backgrounds, including Arabic, Czech, English, French, German, Japanese, Mongolian, Portuguese, Romanian, Russian, and Tigrinya. Three tasks were used to elicit the structures listed in (29). The first two, targeting canonical word order and ADJ topicalisation, although standard in SLA, were adjusted for Italian in order to allow for production of both overt SUBJ and null-SUBJ. Thus, first, a short picture story retelling task involving at least two actors encourages the use of canonical word order with both full referential or pronominal SUBJ and pro-drop. Second, a spot-the-difference task gives speakers the chance to contextualise the different actions depicted in the two drawings by prefacing such phrasal ADJ as *nel disegno a* ('in drawing a') and *nel secondo disegno b* ('in the second drawing b'). Thirdly, for eliciting OBJ topicalisations we adapted Di Biase's 'animal dinner' described in § 2.2, which became a 'birthday party', and a 'trip into the mountains' where various protagonists must contribute a present in the first case and a piece of equipment in the second. From the learners' oral production data set, 512 declaratives sentences could be extracted, about 34 sentences per learner. They are all main clauses with lexical Vs requiring default mapping of thematic roles onto GFs, such as *dare* ('give', a three-argument V), *comprare* ('buy', a two-argument V) or *saltare* ('jump', a one-argument V). In other words, sentences with so-called exceptional Vs and nonbasic V forms such as passives (Pinker 1984) are not considered here, because all of these involve nondefault mapping between a-structure and f-structure, and consequently PT's Lexical Mapping Hypothesis. Nor do we consider copular and presentative sentences, because they are 'nonverbal predicates' (Kroeger 2005: ch. 10).

In (29) we arrange structures vertically from the bottom up with the three groupings coherently reflecting our hypotheses, and learners horizontally from the left according to developmental progress. In general terms, results shows an implicational developmental pattern: all learners who have achieved the XP canonical word order stage have acquired various instantiations of the previous canonical word order stage. Then, likewise, those who have achieved the non-canonical (marked) stage have clearly achieved both previous ones. The fact that the L1 background of our 15 learners is typologically varied makes our evidence even more convincing. This of course does not deny the influence of L1. Transfer does occur, but it is constrained by the procedures the learner can handle (Pienemann, Di Biase, Kawaguchi & Håkansson 2005). As such, it does not affect the developmental sequence, although it may affect the hypothesis space of a given stage, the speed with which it is attained, or the structur-

al preferences it manifests, as shown for RYA. Within this PT-based implicational developmental pattern, a closer look at the figures reveals further interesting patterns.

(29) *Cross-sectional study of syntactic development in L2 Italian based on the Prominence Hypothesis: declaratives (after Bettoni & Di Biase 2011)*

STRUCTURES	GHI	MID	MAR	MAC	CHA	LEI	NAT	RIC	TAN	VAN	SHI	SER	EVA	HEL	RYA
TOP _i Cl _{OBJ.i} -AUX V-to _i SUBJ											-8			4	8
TOP _i Cl _{OBJ.i} -V SUBJ						-4	-7	-12					3	4	8
TOP _{ADJ} SV(O)		4	8	6	9	9	2	4	4	5	2	9	2	4	5
TOP _{ADJ} V(O)			1				2	3		2		1			
V(O)				2			1	3	5	6	9	5	5	6	
SUBJ _{pr} V (O)	2	3	6	5	7	3	8	7	2	2	7	6	3		7
*SUBJ _{ref-THEME} V(O)	4	13	8	11	8	10	13	11		9					
SUBJ _{ref-AGENT} V(O)	13	6	16	16	15	11	8	2	15	10	8	12	9	11	10

In the two OBJ topicalisation rows at the top, the numbers to the left of the vertical line refer to sentences produced with various incorrect V agreements: cf. examples (35)–(37) below.

In the SUBJ_{THEME} row at the canonical word order stage, learners produce an interlanguage form involving nondefault mapping of thematic role (theme) onto GF (SUBJ): cf. example (31), note 17, and § 3.2.

All our learners are beyond the lemma access stage. Even GHI, who is the weakest, produces several sentences with canonical word order, as also found in other studies of L2 Italian using different theoretical frames (e.g., Andorno, Bernini, Giacalone Ramat & Valentini 2003). In first position these sentences include most commonly the agent with default mapping, as in (30), but also – under the pragmatic constraints of our specific task – the theme with nondefault mapping, as in (31). Unable to produce an OBJ_{THEME} topicalisation required by the task we have described above, GHI (like the other less proficient learners) produces instead a clear example of functional assignment specified solely by position. Without labouring the point here, this is a case where the interfacing between the morphological resources and syntactic progress still needs to be worked out in more detail along both the Prominence Hypothesis schedule and the Lexical Mapping Hypothesis one (cf. ch. 1, § 4.3, this volume).

- (30) GHI lui scito [=lasciato] carta identità per scrivere
 he-SUBJ left-PAST PARTICIPLE identity card-OBJ to write
 [he has left (his) identity card (on the counter) in order to write]

- As we move from GHI rightwards in the table, there is significant variation among the learners in the distribution of the different constructions belonging to this canonical word order stage, concerning mainly three points: morphological accuracy of V inflection, the thematic roles in the SVO order, and use of SUBJ vs null SUBJ. With regard to the first point, although a detailed analysis on this set of data is not yet available, it is worth noting that learners to the left of our interlanguage path have poor V inflection – marking mostly aspectual but not grammatical person, unequivocally indicating SUBJ, as in (30)-(31). On the other hand, inflection improves considerably as we move rightward in the table. Secondly, as the figures for SVO show, the weaker learners produce ungrammaticalised topicalisations of themes in first position similar to GHI's in (31). On the other hand, learners to the right clearly avoid them. Thirdly, the table shows how null SUBJ tends to be avoided by the weaker learners, who have unreliable morphological means to assign SUBJ, and then progressively increases rightwards. Here the only exception is RYA – an interesting case, most likely due to the fact that his L1 is English, which requires overt SUBJ. In any case, most learners overuse pronominal SUBJ, as he does in (32), where *lui* ('he') is redundantly repeated three times.

- All our learners except GHI have also reached the XP canonical word order stage, and are able to front ADJ to the left of the canonical sequence. The reader will notice that in (29), as well as the structure with SUBJ, as in (33), we also count the one with null SUBJ, as in (34), both by VAN. This needs an explanation because if there is no SUBJ to compete with TOP, one could think that the evidence for progress to this stage is missing. However, we claim that in a null-SUBJ language, VO is also a canonical string.

- (33) VAN nel disegno B il ragazzo anda nella bicicletta
in the drawing b-ADJ the boy-SUBJ goes-V on the bicycle-OBL
[in drawing B the boy rides a bicycle]

- (34) VAN nel disegno b guardano le borse
 in the drawing b-ADJ look at-V the bags-OBJ
 [in drawing B they look at the bags]

Among our learners, three have progressed also to the highest noncanonical word order stage. Being able to assign functions independently from position, they now have the means to place OBJ and SUBJ in a position other than their canonical one, and mark them with the correct morphological resources. In this cross-sectional study we can now plot the developmental path learners traverse when faced with a task requiring them to begin the sentence with a role other than the agent, and we can discern several phases. In a first phase, represented by GHI, MID, MAR, MAC, CHA, and VAN, learners most often begin the sentence with the theme without grammaticising the topicalisation, as we have seen in (31); at other times however they use canonical order but choose to start with SUBJ as agent, thus ignoring the prompt from the task, as in (35). In a second phase, represented by LEI, NAT, and RIC, learners start to grammaticise the topicalisation of the theme by using a clitic, but the grammaticisation is incomplete in the sense that one, two or all three agreements (SUBJ-V, TOP-clitic, and TOP-clitic-participle) are missing or not unified. For example, in (36) LEI uses the clitic, but neither SUBJ-V nor TOP-clitic agreement is in place; and in (37) RIC mixes features of passivation with those of topicalisation.

- (35) MAR lo stereo.. lo stereo... il cameriere porta il cameriere porta il stereo
 [the stereo, the stereo, the waiter brings the waiter brings the stereo]

- (36) LEI i fragoli la porta le ballerine
 the stawberries-FM.PL it-FM.SG take-3.SG the dancers-3.PL
 [the strawberries [them] take the dancers]

- (37) RIC la torta la porta dai poliziotti
 the cake-FM.SG it-FM.SG takes-3.SG by the policemen-3.PL
 [the cake (it) takes the policemen]

In a third phase, represented by EVA and HEL, learners produce some grammatical and some ungrammatical topicalisations. In the final phase, represented by RYA, all topicalisations are targetlike; and even the third long distance, interphrasal dependency is computed with the correct agreement between the clitic and the participle, as in (38).

- (38) RYA le calze le ha portate l'infermiera
 the socks-FM.PL them-FM.PL have-3.SG taken-FM.PL the nurse-3.SG
 [the socks (them) has taken the nurse]

Although a fine-grained analysis of the increasing morphological accuracy still needs to be done, we consider this sequence an important confirmation that learners may shoot up syntactically in their word order development well before they acquire the full morphological means to unequivocally mark all the GFs in their sentences.

3.2. The Lexical Mapping Hypothesis

Italian Vs lexically requiring nondefault mapping of thematic roles onto GFs are not only numerous in type, but some of them also occur quite frequently as tokens in everyday speech. Among this latter group, we find unaccusative Vs (e.g., *arrivare* 'arrive' and *morire* 'die'), exceptional Vs (e.g., *piacere* 'like' and *mancare* 'be missing'), and nonbasic V forms, such as reflexive Vs (e.g., *lavarsi* 'wash [oneself]' and *arrabbiarsi* 'get angry').

Although as formulas they can be used by L2 learners quite early – see, for example, the ubiquitous *mi piace* ('I like it') –, all these Vs lexically requiring nondefault mapping are acquired – that is, processed online with some grammatical accuracy – quite late, at the final nondefault mapping stage. Given the great variety among them, it is not surprising that they may also be acquired at different times within this last stage, as we have seen in the previous discussion of morphological development. Factors affecting their sequencing include (a) the number of arguments they subcategorise: see, for example, *nascere* ('be born') vs *piacere* ('like'), which require one and two arguments respectively; (b) their information structure; that is, whether, in a pragmatically neutral context, they require noncanonical word order in addition to nondefault mapping: see, for example, passive V forms, which require topical (preverbal) SUBJ, as in (39), vs unergatives, unaccusatives and most exceptional Vs, which prefer focal (postverbal) SUBJ, as in (40), (41) and (42) respectively; (c) the nature and function of their clitic particles in the case of the reflexive V family; that is, whether the clitic particle is purely expletive, as in *pentirsi* ('repent'), which does not have a corresponding nonreflexive form **pentire*, or semantically somewhat more transparent, as in *baciarsi* ('kiss [each other]'), which has a transitive form *baciare* ('kiss [someone]') as well; furthermore, whether the reflexive construction is decausative, as in (43), entails OBJ reduction, as in (44), OBL reduction, as in (45), or no grammatical function reduction but adds a pragmatic modality to the eventuality being described, as in (46).

- (39) la pasta viene servita calda
 SUBJ V
 [pasta is being served hot]
- (40) ha telefonato la nonna
 V SUBJ
 [the grandmother has phoned]
- (41) è nata una bambina
 V SUBJ
 [a baby girl is born]
- (42) ai bambini piace la pizza
 OBL V SUBJ
 [pizza pleases the children = the children love pizza]
- (43) il bus si ferma
 SUBJ V
 [the bus stops]
- (44) si lava
 V
 [(he) washes (himself)]
- (45) si leva la giacca
 V OBJ
 [(he) takes off the jacket]
- (46) si mangia la pizza
 V OBJ
 [(he) eats (his) pizza (loving it)]

In (47) we propose our language specific developmental schedule for Italian based on the Lexical Mapping Hypothesis discussed universally in chapter 1 (cf. § 4.2.2, and ch. 2, § 3.2, for English).

At the default mapping stage, learners may soon lexically saturate all the roles that are semantically relevant to their Vs, provided of course they store the appropriate items in the lexicon. They will, however, be unable to map them onto the appropriate GFs unequivocally, that is by morphological means. So, for Vs lexically requiring a further argument besides those mapped on SUBJ and OBJ, learners will tend to leave it functionally unmarked, as in (48)-(49) by MAR and CHA

(47) *Developmental stages hypothesised for L2 Italian syntax based on the Lexical Mapping Hypothesis*

STAGE	STRUCTURE	EXAMPLE
NONDEFAULT MAPPING	reflexive constructions	<i>Pierino si lava le mani</i> [Pierino washes his hands]
	causatives constructions	<i>il papà lascia guidare la moto a Pierino</i> [daddy lets Pierino drive the motobyke]
	exceptional verbs	<i>a Pierina piace la cioccolata</i> [to Pierina chocolate is pleasing]
	unaccusative verbs	<i>è morto il gatto</i> [the cat has died]
	passive verbs	<i>il pesce blu è mangiato dal pesce verde</i> [the blue fish is eaten by the green fish]
DEFAULT MAPPING AND ADDITIONAL ARGUMENTS	agent/experiencer mapped on SUBJ, patient/theme mapped on OBJ, and other members of the a-structure hierarchy, such as goals and locatives, mapped on GFs other than SUBJ and OBJ	<i>la mamma mette i fichi nel frigo</i> [mum puts the figs into the fridge] <i>la nonna dà un bacio a Pierino</i> [grandma gives a kiss to Pierino]
DEFAULT MAPPING	agent/experiencer mapped on SUBJ and patient/theme mapped on OBJ	<i>il pesce verde mangia il pesce blu</i> [the green fish eats the blue fish]
LEMMA ACCESS	single words formulas	<i>ciao</i> [hello] <i>mi chiamo Pierino</i> [my name is Pierino]

respectively, two of the learners in the corpus mentioned above in § 3.1. Italian, in fact – unlike configurational English – has no OBJ θ , and requires both the beneficiary or source roles to be mapped onto OBL by means of a PP.

- (48) MAR la ragazza è data mh Carlo la chiave
 SUBJ_{AGENT} V OBJ_{GOAL} OBJ_{THEME}
 [the girl has given Carlo the key]
 target la ragazza ha dato a Carlo la chiave
 SUBJ_{AGENT} V OBL_{GOAL} OBJ_{THEME}
- (49) CHA Susanna chiede.. un uomo ahh dove San Pietro
 SUBJ_{AGENT} V OBJ_{THEME} SUBORDINATE CLAUSE
 [Susanna asks a man where Saint Peter (is)]
 target Susanna chiede a un uomo dove è San Pietro
 SUBJ_{AGENT} V OBL_{SOURCE} SUBORDINATE CLAUSE

Likewise, any V types may be used early by beginners. The crucial point, however, is again that in certain cases the outcome will be ungrammatical or pragmatically inappropriate because at this stage Vs are treated as mapping the highest thematic role onto the highest GF in the first linear position. So, besides felicitous outcomes with 'regular' Vs, as in (50), beginners will produce sentences such as that in (51), where MID uses the exceptional V (*piacere*, 'is pleasing') ignoring that in Italian its experiencer role requires to be mapped as TOP_{OBL}.

- | | | | |
|----------|-----------------------------|-----------|-----------------------|
| (50) MAR | la signora | adorare | la:: scarpe |
| | SUBJ _{AGENT} | V | OBJ _{THEME} |
| | [the lady | adores | the shoes] |
| | | | |
| (51) MID | i pagliacci | piacere:: | caramelle |
| | SUBJ _{EXPERIENCER} | V | OBJ _{THEME} |
| | [clowns | like | sweets] |
| target | ai pagliacci | piaccono | le caramelle |
| | OBL _{EXPERIENCER} | V | SUBJ _{THEME} |

At the next stage up, learners begin to mark the GFs of arguments other than SUBJ and OBJ more clearly, and produce sentences such as those in (52)-(53). Here, although the prepositions are lexically inaccurate (*per* and *di* respectively instead of *a*), the very fact of using a PP marks the function of the constituent as OBL, that is as neither SUBJ nor OBJ (the two core GFs of the canonical order). We assume, as explained in chapter 1, § 4.2.2, that this is a necessary intermediate stage before learners acquire the means to map arguments onto GFs in an unequivocal way.

- (52) RIC il marito ha detto [= dato] i fiori per [= a] sua moglie
SUBJ has given OBJ_{THEME} OBL_{BENEFICIARY}
[the husband has given the flower to his wife]
- (53) TAN la receptionista ha chiesto di [= a] Carlo la mh sua carta
di identità
the receptionist-SUBJ V of Carlo-OBL_{SOURCE} his identity card-
OBJ_{THEME}
[the receptionist has asked Carlo for his identity card]

At the final nondefault mapping stage learners will then be able to assign arguments to any thematic roles as required by their lexical entries, and hence assign prominence to thematic roles low in the hierarchy by mapping them onto SUBJ, as in (54), or demote those that are high up by mapping them, for exam-

ple, onto OBL or ADJ, or even suppress them, as in (55). These latter two examples are fictitious.

(54) al ministro mancano i fondi
 OBJ_θEXPERIENCER V SUBJ_{THEME}
 [the minister is lacking the funds]

(55) il presidente è stato fischiato
 SUBJ_{THEME} V
 [the president has been booed]

Empirical work testing our Lexical Mapping Hypothesis for L2 Italian has just begun. Needless to say parts of our schedule in (47) are not yet supported by actual data. We mention here three studies in chronological order, which are heterogeneous for the variety of Vs and constructions they deal with. Bettoni, Di Biase & Nuzzo (2009) deal with nondefault mapping requirements from the point of view of the acquisition of postverbal SUBJ, and show that, within the last stage of development, unergative and unaccusative Vs are acquired with their required mapping before exceptional Vs. Nuzzo (2012) deals with the acquisition of passives, and clearly shows how long and complex the path toward their full acquisition can be, despite the fact that structurally they require canonical word order. The point Nuzzo makes is that even learners who can produce SVO sentences with a high degree of grammatical accuracy – as long as default mapping is required – may fail to do so when mapping requirements are nondefault. The explanation, we assume, is that they fail to integrate pragmatic and structural information in order to give their sentences a suitable perspective so as to guide the listeners' attention according to their communicative intentions. The third study testing the Lexical Mapping Hypothesis is by Bettoni & Fratter (2013), and deals with Italian reflexive V forms.

4. Conclusion

As we have seen, Italian is an interesting language for PT in so far as it presents different challenges compared to English, chiefly posed by its rich obligatory morphology licensing more flexible options in word order. Since the seminal work on L2 Italian by Di Biase & Kawaguchi in 2002, the analyses mentioned in this chapter have not only added to our knowledge of the development of this language but they have also paved the way for other nonconfigurational languages. More discussion and results will be presented in Bettoni & Ginelli's chapter 8 on the syntactic development of content questions.

This chapter also takes up the challenge of identifying the sources of sequencing regularities noted within a particular stage, and indicates both how they can be conceptualised by identifying soft (inrastage) barriers as compared to hard (interstage) barriers. Furthermore, this chapter offers a key for predicting their sequencing. The quality and quantity of features to be processed at any given moment and how specific DFs are involved will offer a principled key for going beyond the more traditional PT predictions and offering a more detailed map for possible language intervention. We have also indicated how looking at the variable behaviour of learners, i.e., at their accuracy in a principled way, rather than ignoring it as soon as one can establish emergence of one particular structure, can be quite revealing of more fine-grained developmental patterns. In any case more work needs to be done, foremost in two directions, in our opinion.

First, with regard to obligatory structures, in a language with such a syncretic morphology as Italian, the challenges posed by the emergence criteria need more attention, as Pallotti (2007) lucidly states. But there are other crucial problems to disentangle along the long path from emergence to acquisition. For example, in strict LFG terms, the two sentences in (56) would both belong to the category procedure stage (with no information exchange outside the VP), because LFG sees clitic pronouns as verbal morphology rather than separate words.

- (56) a. glieli dò oggi
 he-DAT they-ACC.PL.MASC give-1.SG today
 [I give them to him today]
- b. glieli ho dati ieri
 he-DAT they-ACC.PL.MASC have-1.SG given-PL.MASC yesterday
 [I have given them to him today]

Yet in order to produce them, learners will need to identify the GFs of all three argument roles of the V *dare* ('give'), and formally distinguish between OBJ and OBL_{DAT} – not to mention the added complexity of marking number and gender on the clitic pronoun and on the past participle. Could we then not assume that learners will proceed from sentences with referential nouns or full pronouns to those with clitics, as in (57a), and from just one clitic, as in (57b), to both of them, as in (56 above) only after they have reached the stage of default mapping and additional arguments in the Lexical Mapping Hypothesis hierarchy?

- (57) a. dò i libri a lui oggi
 give-1.SG the books to him today
 [I give the books to him today]

b. li	dò	a lui	oggi
they-ACC.PL.MASC	give-1.SG	to him	today
[I give them to him today]			

Could we not also assume that learners will produce accurate sentences first with SUBJ morphology, then with clitic OBJ markers, and finally clitic with OBL_{DAT} ones? See, in this regard, the stage of default mapping and additional arguments in (47), § 3.2. The rich nature of categorial markings in Italian offers a good area for testing hypotheses of morphological development across soft barriers within stages, and its interfaces with syntactic development.

Secondly, in terms of the acquisition of optional or alternative syntactic constructions, PT can be quite precise not only on structural optionality dealt with by the Prominence Hypothesis, but also on the lexical-semantic requirements dealt with by the Lexical Mapping Hypothesis. In this direction, Italian has a lot of ground to offer to both theory development and empirical testing.

The development of Japanese as a second language

Satomi Kawaguchi

University of Western Sydney

1. Introduction

Japanese has an important role to play for testing and developing PT's hypotheses because of its typological characteristics. With regard to the configurationality spectrum which has languages like English and Warlpiri at opposite ends (cf. § 2.2, ch. 1, this volume), Japanese is located far from configurational English and closer towards nonconfigurational Warlpiri. Morphologically, Japanese is rich in verbal inflections, but unlike Italian its morphological organisation is agglutinating rather than fusional, and also unlike Italian its grammatical relations are marked on nominal constituents rather than on the verb, which makes Japanese a dependent-marking language. In this latter sense it is similar to Russian and Serbian (cf. chh. 5 and 6 respectively). Syntactically, Japanese is an SOV, head-last language which allows great freedom in the ordering of nominal constituents, as long as V is sentence-final. For a fuller description of the Japanese language, written in English, readers are referred to Kuno (1973) and Shibatani (1990). For a shorter survey, see Shibatani (2011).

This chapter represents a new development in the studies on L2 Japanese in so far as, first, it updates the schedules for morphological and syntactic development outlined in earlier work according to the changes to PT proposed by Bettoni & Di Biase in chapter 1 of this volume, and, second, in doing so, it explores the boundaries of PT-based hypotheses on the acquisition of Japanese syntax. Work on the acquisition of Japanese nominal and verbal morphology in the PT framework began with Doi & Yoshioka (1987, 1990), Huter (1996), Kawaguchi (2000, 2002), and Di Biase & Kawaguchi (2002). Then, Japanese L2 studies contributed substantially to the extended scope of PT. For example, Kawaguchi's (2005) longitudinal study was the first to test the two new hypotheses known as "Extended PT" (i.e., the Topic Hypothesis and the Lexical Mapping Hypothesis proposed in Pienemann, Di Biase & Kawaguchi 2005). Furthermore, Kawaguchi (2007) reports on the acquisition of structures involving nonbasic V forms such as passives, causatives and benefactives, as well as exceptional Vs such as unaccusatives, all of which lexically require nondefault mapping – I use here Pinker's

(1984) expressions ‘nonbasic’ for V forms and ‘exceptional’ for V types introduced in ch. 1, § 4.2.2, this volume. Further, Kawaguchi’s (2009a) cross-sectional study involving 24 intermediate-advanced university learners presents a detailed analysis of the acquisition of causative constructions in relation to the Lexical Mapping Hypothesis. So far, this hypothesis has been applied also to other languages such as English (Kawaguchi 2013; Keatinge & Keßler 2009; Wang 2009) and Italian (Bettoni & Fratter 2013; Nuzzo 2012), but Japanese remains the only one for which it has been shown how to treat causative constructions within the PT framework. Kawaguchi’s series of studies deal with Japanese as a foreign language in an instructional setting involving university students in Australia, but her results find support in other contexts, such as those on child L2 acquisition of Japanese in a naturalistic environment (Iwasaki 2008), adult language acquisition of L2 Japanese in an intensive course (Iwasaki 2013), and significantly bilingual L1 acquisition in Japanese-English (Itani-Adams 2007, 2008, 2009, 2011, 2013). The current state of PT as applied to L2 Japanese involves in particular the development of syntax and a reanalysis of the data from Kawaguchi (2010) with regard to the Prominence Hypothesis, and Kawaguchi (2005) with regards to the Lexical Mapping Hypothesis. About the latter hypothesis, Japanese as an L2 is the only language besides English (cf. § 3.2, ch. 2, this volume) that tests empirically the intermediate stage of default mapping plus additional arguments, formally introduced by Bettoni & Di Biase in chapter 1, § 4.2.2.

In the following sections each developmental schedule is preceded by a sketch of the main characteristics of its respective Japanese target structures, and followed by the pertinent empirical data.

2. Morphological development

Japanese is an agglutinating language, characterised as predominantly suffixing rather than prefixing in its inflectional morphology (Dryer 2013). With regard to verbal morphology, various affixes added to the stem of a V provide meanings such as tense, aspect, mood, politeness, and polarity. In English a notion similar to agglutinative morphology is the addition of a morpheme (e.g., past *-ed*, progressive *-ing*) to the lexical V, of a separate AUX (e.g., *can*, *may*, *must*), or a separate main V (e.g., *cause*, *want* and *seem* for causative, desiderative and evidential constructions respectively). In Japanese these morphemes are attached immediately after the V stem. When more than two suffixes are sequenced, their relative order is generally fixed. For example, the passive suffix must precede the tense suffix, as in (1a), which is grammatical, whereas (1b) is ungrammatical.

- (1) a. *tabe-rare-ta*
 eat-PASSIVE-PAST
 [(it) was eaten]
- b. **tabe-ta-rare*
 [eat-PAST-PASSIVE]

Because Japanese morphology is agglutinating it is relatively easier to segment for the learner than Italian fusional morphology. This can be seen in the example (1a) above where the verb stem (*tabe-* ‘eat’), the passive auxiliary verb (*-rare*), and the verbal morphological inflection expressing past tense (*-ta*) are all uniquely identifiable. On the other hand, complex morphophonological processes are involved in affixation. The basic Japanese syllable structure is CV whereby phonological rules intervene to avoid consonant clusters and vowel clusters in the affixation process. For example, the initial consonant of the suffix is elided when a consonant-initial verbal morpheme, such as the present tense suffix *-ru*, is added to the consonant-final stem as in (2a). Similarly the initial vowel of the suffix is elided when a vowel initial suffix, such as the negative suffix *-anai*, is added to the vowel-final verb stem as in (2b) (Shibatani 1987). However, these morphophonological processes lie outside the current scope of PT, so they will not be mentioned further here.

- (2) a. *kak-u*
 write-PRES
 [write]
- b. *mi-nai*
 see-NEG
 [do/does not see]

With regard to nominal morphology, Japanese displays both structural and semantic case markers. Consistently with the characteristics of a head-last language, these are all postpositional particles. There are four structural markers: *-ga* for nominative case, *-o* for accusative case, *-ni* for dative case, and *-no* for genitive case, as shown in (3).

- (3) Harumi-ga imooto-no tegami-o Taro-ni age-ta
 Harumi-NOM sister-GEN letter-ACC Taro-DAT give-PAST
 [Harumi gave (her) sister’s letter to Taro]

Then there are several semantic case markers, such as *-de* and *-kara*.

- (4) a. Taro-ga pen-de tegami-o kai-ta
 Taro-NOM pen-INSTR letter-ACC wrote-PAST
 [Taro wrote the letter with a pen]
- b. Harumi-ga Tokyo-kara ki-ta
 Harumi-NOM Tokyo-ABL came-PAST
 [Harumi came from Tokyo]

Other Japanese nominal morphemes include numeral classifiers, which are common in many East Asian languages such as Chinese and Korean. In Japanese, classifiers always follow numeral modifiers, and are chosen according to what type of object is being counted. So, counting animals, or books, requires different classifiers, as (5a-b) show.¹

- (5) a. ni-hiki-no inu
 two-CL-GEN dog
 [two dogs]
- b. ni-satsu-no hon
 two-CL-GEN book
 [two books]

A further, and highly characteristic, morpheme of Japanese nominals is the discourse marker *-wa* indicating primarily the sentence TOP, as in (6), but also the sentence FOC. We will discuss this marker in § 3 below when dealing with syntax.

- (6) sensei-wa kohii-o nomi-masu
 teacher-TOP coffee-acc drink-pol
 [teacher drinks coffee]

Finally, just as Japanese Vs do not mark SUBJ information regarding NUM, so Ns do not mark NUM or GEN distinctions – although there are the plural N suffixes *-tachi* and *-ra*, as in *kodomo-tachi* (child-PL, “children”) and *boku-ra* (1.PL, “we”), but these are not productively used and can only be attached to a limited number of lexical items. Likewise, just as there is no SUBJ V agreement, so there is no agreement between adjectives or numerals and nouns.

1 Japanese nominal classifiers have not been investigated in PT studies – a good area to explore.

The developmental hypotheses for Japanese verbal morphology proposed in Di Biase & Kawaguchi (2002) are shown in (7). Nominal morphology is not dealt with in this chapter, except when it is deployed in syntax for marking GFs (cf. § 3).

(7) *Developmental stages hypothesised for L2 Japanese morphology (after Di Biase & Kawaguchi 2002)*

PROCEDURE	MORPHOLOGICAL OUTCOME/STAGE	STRUCTURE	EXAMPLE
S-BAR PROCEDURE	INTERCLAUSAL MORPHOLOGY	GEN on SUBJ in N-modifying clause	<i>Kumiko-no sotugyosita daigaku-ga setagaya-ni arimasu</i> [the university Kumiko graduated from is in Setagaya]
S-PROCEDURE	INTERPHRASAL MORPHOLOGY	nondefault case marking in constructions such as passive, causative, benefactive	<i>sakana-ga neko-ni tabe-rare-masita</i> [the fish was eaten by the cat]
PHRASAL PROCEDURE	PHRASAL MORPHOLOGY	N-GEN N	<i>inu-no namae</i> [dog's name]
		V-te-V (V-COMP V)	<i>hanasi-te mimasu</i> [I try speaking (to them)] <i>mi-te imasu</i> [I am watching]
CATEGORY PROCEDURE	LEXICAL MORPHOLOGY	case marking on N: e.g., NOM-ACC alternation	<i>sensei-ga / sensei-o</i> [teacher]
		V inflection: e.g., present-past alternation	<i>tabe-masu / tabe-masita</i> [eat / ate]
LEMMA ACCESS	INVARIANT FORMS	single words formulas	<i>oisii</i> [delicious] <i>arigatoo</i> [thank you]

The developmental stages which can be hypothesised for Japanese verbal morphology using the framework of PT are described below. At the very beginning, like learners of any other language, learners of Japanese also cannot activate any language-specific procedure, and are thus able to produce only invariant words such as *oisii* ('delicious'), or fixed expressions such as *arigatoo* ('thank you').

As soon as learners are able to activate the category procedure, lexical variation results in some V inflection. In L2 Japanese the most common alternation is between present tense and past tense, as in (8a-b). Furthermore – although, as noted above, Japanese Vs do not inflect for person or number of SUBJ – other

Vstem–affix combinations can indicate the acquisition of lexical operations. Among them, we find Vstem-NEG, as in (8c).

- (8) a. *tabe-masu*
 eat-POL
 [(I/you/he/she, etc.) eat]
- b. *tabe-masi-ta*
 eat-POL-PAST
 [(I/you/he/she, etc.) ate]
- c. *iki-mas-en*
 go-POL-NEG
 [(I/you/he/she, etc.) do/does not go]

Nominal morphology can also be activated at the stage of categorial procedure, and post-nominal particles start to appear. However, the categorial procedure does not permit the recognition of grammatical functions of NP in a sentence. Therefore, the learner identifies nominal morpheme *-ga* (NOM) to mark agent-like NP, and *-o* (ACC) to mark patient- or theme-like NP, which are likely associated with the initial and pre-verbal position respectively. These markers will be taken up again in § 3.2 when dealing with the Lexical Mapping Hypothesis.

Phrasal morphology emerges at the next stage. A Japanese V can combine with another V. In such case, the first needs to be a gerund (*-te*) (Kageyama 1999), where *-te* functions as complementiser (COMP). *V-te V* is an example of phrasal procedure because information exchange is required between two Vs in terms of the ‘combinatoric TYPE’ feature whereby the main lexical V (head element in VP) takes gerundive form in order to combine with the auxiliary V (Sells 1995; 1999). Sells explains that the lexical feature TYPE holds crucial information for verbal projection (i.e., phrasal syntax). In *V-te V* construction, TYPE of *V-te* is *V-sis(ter)*. This means that *V-te* has to take another V as its sister. Thus, the construction *V-te V* requires information unification between two Vs in terms of TYPE, which means that its production requires the phrasal procedure. The Vstem-complementiser V construction (*V-te V*) is one of the ways in which two Vs can combine to add progressive aspect (*V-te imasu*, ‘be V-ing’), as in (9a), trial (*V-te mimasu*, ‘try V-ing’), as in (9b), and request (*V-te kudasai*, ‘please V’), as in (9c).

- (9) a. *hasit-te i-masu*
 run-COMP PROG-POL
 [(I/you/he/she, etc.) is running]

- b. hasit-te mi-masu
 run-COMP try-POL
 [(I/you/he/she, etc.) try running]
- c. hasit-te kudasai
 run-COMP REQUEST
 [please run]

At the next stage, Japanese requires the activation of the S-procedure for marking GFs of NPs in sentences involving nondefault mapping between a-structure and f-structure, such as passive, causative and benefactive constructions. Thus, the morphological operations at the S-procedure stage will be discussed in § 3.2, which illustrates syntactic development at the level of the sentence, where nondefault case marking of Ns interacts with the predicate of the sentence.

At the S-BAR procedure stage, PT hypothesises one operation in Japanese morphology, which is the marking of the GEN case on SUBJ in a N-modifying clause. SUBJ in Japanese is usually marked as NOM, but in a relative clause it can be marked as NOM or GEN, as exemplified in (10) – a phenomenon termed *Ga/No Conversion* by Harada (1971). This requires the learner to be able to distinguish between the subordinate clause and main clause.

- (10) Kumiko-ga/no osie-ta gakusei-o mi-masi-ta
 Kumiko-NOM/GEN teach-PAST student-ACC see-POL-PAST
 [(I) saw a student who Kumiko taught]

Evidence in support of the morphological schedule outlined in (6) comes mainly from two longitudinal studies and two cross-sectional studies involving 28 adult learners of Japanese as an L2. Further details of these four studies can be found in Kawaguchi (2010). The data in (11) illustrates the results of one of these studies – a three year longitudinal study of a learner codenamed Lou – used here to exemplify the morphological development observed over the fuller data set. Lou is a young woman, native speaker of English studying L2 Japanese at an Australian university. Her data was collected at 12 points in time over her three years as an undergraduate, that is, twice a semester over six semesters, yielding a total of about 3,200 token words and 1,345 type words.

Note that the data in (11) includes Lou's development for verbal morphology up to the VP procedure stage. In fact, as we have just mentioned, structures requiring the S-procedure involve nondefault mapping and they will be discussed in § 3.2. Further up, the hypothesised Japanese L2 morphology at the S-BAR stage has yet to be empirically tested, just as is the case of L2 English and L2 Italian.

(11) *Lou's morphological development (after Kawaguchi 2010)*

STAGE	STRUCTURE	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12
VP PROCED.	V-TE V												
	V-te plus other V									+1		+3	+1
	V-te REQ					+2				+1			
	V-te BEN					+2			+1	+1	+1		+5
	V-te ASP				+8	+8	+2	-1	+1	+4	+2	+4	+2
CATEGORY PROCED.	VERBAL INFLECTION												
	Vstem-PLAIN-PRES/PAST-NEG											+1	+1
	Vstem-POL-PAST-NEG								+2	+4		+1	+5
	Vstem-POL-PRES-NEG				+2	+1	+1	+1	+1	+2	+4	+3	+3
	Vstem-PLAIN-PAST					+2	+1		+4	+2	+4	+7	+9
	Vstem-POL-PAST			+12	+13	+21	+14	+1	+11	+30	+8	+21	+16
	Vstem-POL-PRES-Q	+2			+2								
	Vstem-PLAIN-PRES	+1				+1	+1	+1		+1	+2	+5	+6
	Vstem-POL-PRES	+7	+19		+8	+1	+5	+5	+24	+12	+13	+15	+14
+ = supplied in obligatory context; - = not supplied in obligatory context; empty cell = no context													
S U M M A R Y													
VP PROCED.	V-TE V	-	-	-	+	+	+	+	+	+	+	+	+
CATEGORY PROCED.	VERBAL INFLECTION	+	+	+	+	+	+	+	+	+	+	+	+
+ = emerged; - = not emerged													

As (11) shows, at t1 Lou already produces several inflectional forms of V: the POL form 7 times, the POL-Q form twice, and the PRES form once. In this session, she produces the POL inflectional form with three different Vs. Furthermore, the Vs *hanasu* ('talk') and *iru* ('exist') are realised in two different inflectional forms, namely, *hanasi-masu* (talk-POL) and *hanas-u* (talk-PRES), as well as *i-masu* (exist-POL) and *i-masu ka* (exist-POL Q) respectively. Thus, at t1, with both lexical and formal variation, the emergence criterion is satisfied. Thereafter Lou continuously shows V inflection throughout the 3-year longitudinal study.

The phrasal procedure stage is reached at t4, when Lou produces the V-te V structure 8 times, as in (12), where *-de* is an allomorph of *-te*; thereafter she uses this form consistently with a variety of lexical items.

- (12) t4 futari kodomo-ga ason-de i-masu
 two children-NOM play-COMP PROG-POL
 [two children are playing]

Thus Lou's developmental sequence for verbal morphology shows an implicational pattern, indicating that the hypothesised hierarchy is supported with 1.0 scalability. This same pattern is confirmed for child L2 Japanese JSL (Iwasaki 2008) and for adult L2 Japanese (Iwasaki 2013), as well as for English-Japanese bilingual L1 acquisition (Itani-Adams 2013).

3. Syntactic development

As mentioned in § 1, Japanese is an SOV, head-last language allowing great freedom in the order of nominal constituents as long as V is sentence-final. Moreover, Japanese is interesting to PT for its rich range of V structures (such as passives, benefactives, causatives) requiring nondefault mapping of a-structure onto f-structure. Accordingly, after the very first stage of single words and formulas, PT organises the learning constraints for syntax around the two hypotheses sketched out first universally in chapter 1, § 4.2 in this volume, and then for L2 English and L2 Italian in chapters 2 and 3 respectively. In fact, unlike Pienemann, Di Biase & Kawaguchi in 2005, I now believe with Bettoni & Di Biase that it is crucial to separate word order issues from mapping issues at the initial stage of L2 syntactic development. For this reason, I will now illustrate the Prominence Hypothesis in § 3.1 and then the Lexical Mapping Hypothesis in § 3.2 for L2 Japanese including their respective initial states, thus making Pienemann, Di Biase & Kawaguchi's (2005) Unmarked Alignment Hypothesis redundant.

3.1. *The Prominence Hypothesis*

As a dependent-marking language, Japanese deploys case markers on Ns to indicate their grammatical relations with V (i.e., GFs). The marker *-wa* indicates discourse relations (i.e., DFs) of topic, focus and contrast within the sentence (cf. § 2.2, ch. 1, in this volume).

With regard to GFs, by default the NOM case marker *-ga* is associated with the GF SUBJ, the ACC marker *-o* with OBJ, and the DAT marker *-ni* with OBL_{RECIPIENT/GOAL}. Being thus marked by case particles, GFs may be placed variably in c-structure – a phenomenon called 'scrambling' in Japanese. The examples below show scrambling in contrast to canonical order, where (13a) exhibits canonical order – SOV in Japanese – and (13b) a scrambled order in which OBJ is linked to the prominent initial position.

- (13) a. Harumi-ga Taro-ni tegami-o yon-de age-ta
 Harumi-NOM.SUBJ Taro-OBL_{RECIPIENT} letter-ACC.OBJ read-COMP BEN-PAST
 [Harumi read a/the letter to Taro]

- b. tegami-o Harumi-ga Taro-ni yon-de age-ta
 letter-ACC.OBJ Harumi-NOM.SUBJ Taro-OBL_{RECIPIENT} read-COMP BEN-PAST
 [Harumi read **a/the letter** to Taro]

In theory, more permutations in c-structure are possible in Japanese as long as V is in final position. Scrambling brings a semantic effect of mild emphasis on the fronted constituent among other effects such as aiding interpretation of antecedent-pronoun relationship (Shibatani 1990: 261). When speakers wish to add prominence to their GFs, they use the discourse marker *-wa* in preference to position (Tsujimura 1996).

So, with regard to the DFs, Japanese (like Korean and a few other languages) has a special marker (*-wa*) for NPs to mark discourse prominence relations within the sentence. Most often *-wa* indicates TOP as it “singles out an accompanying noun as the topic of the sentence” (Tsujimura 1996: 135), but *-wa* marked phrases should not be defined just as ‘topic’ because this marker may also indicate focus and contrast (Vermeulen 2009). Notice here, importantly, that such morphological marking of DFs does not require that the *-wa* marked function be in sentence initial position, as shown in (18) and discussed below. Topic, Focus and Contrast are notions of information structure that interact with syntax (e.g., Frey 2004), and it has been recently argued that contrastive focus should be analysed as a composition of the notions [focus] and [contrast] (Vermeulen 2009: 335). This issue is clearly beyond the scope of this chapter, but see chapter 1, § 4.2.1, this volume where information structure features such as [\pm Prominent] and [\pm New] from Choi (2001) are used to constrain the TOP-FOC and their relative prominence. In this connection it may be worth noticing that, interestingly, *-wa* is called a ‘TOP marker’ in English, whereas in Japanese it is called *toritate jyoshi*, which can be translated as ‘bringing up into prominence’ – surely a more suitable descriptor given the use PT makes of it in this volume (cf. the Prominence Hypothesis proposed in (34)–(35), ch. 1, § 4.2.1.).

Since every DF must be linked to a GF, *-wa* can overlay a range of GFs. For example, *-wa* can bring into prominence the core GFs, namely SUBJ² or OBJ, as in (14) and (15) respectively, a nonargument GF, namely ADJ, as in (16), just as it can mark prominence on any other argument function.

2 The reader may recall from § 2.2, ch. 1, this volume, that most languages do not distinguish explicitly TOP from SUBJ, which leads to an interpretation of SUBJ as TOP by default.

- (14) *watasi-wa kinoo kono tegami-o kai-ta*
 I-TOP_{SUBJ} yesterday-ADJ this letter-ACC.OBJ write-PAST
 [I wrote this letter yesterday]
- (15) *kono tegami-wa kinoo watasi-ga kai-ta*
 this letter-TOP_{OBJ} yesterday-ADJ I-NOM write-PAST
 [this letter yesterday I wrote]
- (16) *kinoo-wa watasi-ga kono tegami-o kai-ta*
 yesterday-TOP_{ADJ} I-NOM.SUBJ this letter-ACC.OBJ write-PAST
 [yesterday I wrote this letter]

Notice also that when TOP marks the core arguments SUBJ or OBJ, *-wa* replaces the NOM or ACC case markers altogether, which supports LFG's differentiation of core from noncore GFs. In fact, in noncore GFs the prominence marker *-wa* is added to the case-marked nominals, as the grammaticality of (17) can attest, in which a noncore DAT OBL already marked with *-ni* is attributed prominence in the sentence by the addition of the *-wa* marker.

- (17) *Harumi-ni-wa kinoo watasi-ga kono tegami-o kai-ta*
 Harumi-DAT.OBL-TOP yesterday-ADJ I-NOM.SUBJ this letter-ACC.OBJ write-PAST
 [to Harumi I wrote this letter yesterday]

What is the difference, then, between the sentences in (15) through to (17) and their possible equivalent without the *-wa* marker? That is, with *kinoo-wa* simply as *kinoo* in (14) or *watasi-wa* replaced by *watasi-ga* in (15), *tegami-wa* by *tegami-o* in (16), and *Harumi-ni-wa* simply by *Harumi-ni* in (17)? The answer involves prominence. Let us consider (14): the sentence with *watasi-wa* would answer the question 'what did you do?', where the topic-comment structure indicates that TOP_{SUBJ} has been established previously. On the other hand, the equivalent sentence with *watasi-ga* would answer the question 'what happened?', where the SUBJ-predicate structure is 'event-reporting' and the whole sentence is 'new information'. Hence, when speakers do express TOP_{SUBJ} they mean to identify it unambiguously. Likewise, the sentence with *tegami-wa* would answer a question about the letter, and its equivalent with *tegami-o* a more general one about the whole event. With these examples we can see clearly that, although there is a universal preference for TOP to be encoded, as Levelt (1989: 260) says, in a "syntactically prominent position",³

3 Levelt goes on to explain that "syntactically prominent [...] can mean that the topic is encoded as a grammatical subject [...]. It can, alternatively, mean that the topic will be encoded early in the sentence, whether or not in the role of subject."

position in c-structure is not the only defining criterion in Japanese. Indeed, on the one hand, morphology may support prominence assigned by sentence-initial position, as confirmed by the grammaticality of (17), where the OBJ_{DAT} *Harumi-ni* is marked as TOP with *-wa*, so that we know that topical prominence is assigned to this participant; such topicalised NPs usually does occupy a clause initial position in Japanese. On the other hand, in the competition between syntactic position and morphological marking of TOP, it is the morphology that wins out in establishing topicality, as (18) shows. In this example, TOP is not the initial ADJ *san-nen mae-ni* ('three years ago'), but, unambiguously in Japanese, the NP *watasi-wa* (I-TOP). So the morphologically *-wa* marked GF, and not the initial GF, win out as the topic of the sentence.

- (18) *san-nen mae-ni* *watasi-wa* *nihon-ni* *iki-masi-ta*
 three-year ago-at-ADJ I-TOP Japan-OBL_{LOC} go-POL-PAST
 [three years ago I went to Japan]

As mentioned above, *-wa* may mark FOC, as well as TOP, which explains why the Japanese expression for it (*toritate jyoshi*, 'bringing up into prominence') is more appropriate than the English gloss, which conflates the two into TOP. In fact in Japanese it is quite possible to mark more than one NP with *-wa*. In this case, the initial *-wa* is TOP and the other is commonly assumed to be FOC (Kuno 1973: 30ff), as in (19). In this example we can see again how morphology and syntax interact in marking GFs, and how position becomes relevant (first TOP, then FOC – as confirmed by Peter Sell, personal communication, 2005) because morphology is opaque, in so far as both TOP and FOC are marked with *-wa*.

- (19) *watasi-no e-wa* *onnanoko-wa* *akai doresu-o* *ki-te i-masu*
 I-GEN picture-TOP_{ADJ} girl-FOC_{SUBJ} red dress-ACC wear-COMP PROG-POL
 [in my picture it is the girl who is wearing a red dress]

Significantly, in discussing the Japanese *-wa* phrase, Shibatani (1990: 277) highlights a difference between forms that "do represent a true experiential judgement structure", such as the *-wa* marked SUBJ and OBJ in (14) and (15), and those that do not, such as the *-wa* marked postpositional phrases and adverbials in (16) through to (19). The latter are not true experiential judgement topics, but in his view they exploit the highlighting function of the *-wa* marker of true topics to set themselves apart from the rest of the sentence. Hence Shibatani (1990: 277) concludes that "(t)he adverbial topic, including a postpositional one, is basically a stylistic topic." This view would have definitely caused problems with the original Topic Hypothesis (Pienemann, Di Biase & Kawaguchi 2005), but it is compatible with the current Prominence Hypothesis (cf. ch. 1 § 4.2.1, this volume), which

embraces both TOP and FOC with their information structure features [\pm Prominent] and [\pm New] (Choi 2001) to ensure a clear demarcation between DFs and prominence assignment.

This brief discussion of the *-wa* particle does not do justice to the complexities of its uses in Japanese native speakers' speech. As a matter of fact, the whole issue of the topicality vs focality of *-wa* is still controversial – involving as it does such areas as topic continuity, nominal ellipsis and noncanonical case-marking. A parallel case can be made for the function of 'scrambling' in word order to which, controversially, is also attributed no obvious function (Hayashi, Tomlin & Yokota 2002), or in any case an "inconclusive" one (Kuno 1973), or, alternatively, one of "mild emphasis" (Shibatani 1990).

Given these complexities, it is not surprising that learners will only gradually acquire the means of expressing GFs and assigning prominence to the NPs in their Japanese sentences. Here I will first illustrate in (20) how learners of L2 Japanese are hypothesised to acquire the encoding of TOP in declaratives, and then report in (21) the results with respect to declaratives using the same longitudinal study that provides the data for morphological development in § 2 (cf. Kawaguchi 2010). In (21), all the sentences involving lexical verbs with at least one argument are analysed. Exceptions are copula and presentational verbs such as *naru* ('become') and *aru/iru* ('there is-') and verbs taking XCOMP such as *omou* ('think') and *iu* ('say'), which is consistent with the analysis in English and Italian respectively in chapters 2 and 3, this volume. Further evidence supporting the hypothesised schedule in (20) comes from two other studies: Kawaguchi (2005) and Itani-Adams (2009).

(20) *Developmental stages hypothesised for L2 Japanese syntax based on the Prominence Hypothesis: declaratives (after Kawaguchi 2005)*

STAGE	STRUCTURE	EXAMPLE
NONCANONICAL WORD ORDER	TOP _{OBJ} SUBJ V	<i>kono tegami-wa Tanaka-san-ga kaita</i> [this letter Mr Tanaka wrote]
XP _{TOP} CANONICAL WORD ORDER	TOP _{ADJ} SOV	<i>kinoo-wa Tanaka-san-ga kono tegami-o kaita</i> [yesterday Mr Tanaka wrote this letter]
CANONICAL WORD ORDER	TOP _{SUBJ} OV	<i>Kaori-san-wa sukaato-o kaimasita</i> [(Miss) Kaori bought a skirt]
LEMMA ACCESS	single words formulas	<i>oisii</i> [delicious] <i>arigatoo</i> [thank you]

(21) *Lou's syntactic development based on the Prominence Hypothesis: declaratives (re-analysis of Kawaguchi 2010)*

STAGE	STRUCTURE	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	
NONCANONICITY IN MARKING NOMINALS	SUBJ _{TOP-WA} OBJ _{FOC-WA} V										1		1	
	OBJ _{TOP-WA} (S)V				1									
XP _{TOP} CANONICAL WORD ORDER	ADJ _{TOP-WA} S(O)V/(S)OV			2*	3				1		2	5	2	
	ADJ S(O)V		5*		3	3	2	1	1	4	6	6	7	
CANONICAL WORD ORDER	SUBJ _{TOP-WA} (O)V				3	5	7	3		7	9	7	2	9
	S(O)V / (S)OV		1*	4*	7	5	8	9	5	10	22	11	26	12
* SUBJ is not expressed														

* SUBJ is not expressed

The Prominence Hypothesis predicts that, after the initial stage of single words and formulas, the learner can produce canonical word order, which cognitively speaking represents the most harmonious linking between c-structure and f-structure. In Japanese, canonical word order yields the sequence SOV or TOP_{SUBJ} OV as in (22). Because Japanese allows nominal ellipsis, canonical word order also includes V-last structures with at least one core argument, either SUBJ or OBJ, as in (22a) and (22b) respectively.

(22) sensei-wa koohii-o nomi-masu
 teacher-TOP_{SUBJ} coffee-ACC.OBJ drink-POL
 [teacher drinks coffee]

(23) a. sensei-wa nomi-masu
 teacher-TOP_{SUBJ} drink-POL
 [teacher drinks (it)]

b. koohii-o nomi-masu
 coffee-ACC.OBJ drink-POL
 [(she) drinks coffee]

In Lou's data, SUBJ is initially elipted in t1 and t2. Thus she uses only one argument (i.e., OBJ) when required in these two sessions. SUBJ marking with *-wa* emerges at t3 and is then consistently produced throughout the longitudinal study.

The next stage, the XP_{DF} canonical word order stage, is characterised by the learner's ability to place ADJ in sentence-initial position, so as to express contextual information (time, place of the event, etc.). This operation, which does not disturb canonical order, triggers however a disengagement of SUBJ from its canoni-

cal first position in the clause. ADJ may then be TOP – conceding that it may be a stylistic TOP, in Shibatani's (1990: 277) term – and we can clearly see whether learners are able to disentangle the TOP marker *-wa* from its initial exclusive association with SUBJ. So, in addition to being able to add a constituent to canonical order, the learner can now assign to it the TOP function by marking it with *-wa*, as in (19) above.

In Lou's data, ADJ_{TOP} marked with *-wa* first appears twice in t3. However, both sentences do not express SUBJ overtly, so that she starts producing the ADJ_{TOP} S(O)V structure with an overt distinction between SUBJ and TOP only at t4, as shown in (24). Marking thus either TOP or SUBJ by *-wa*, Lou now knows that the sentence initial element can be ADJ_{TOP}, that is, a GF other than SUBJ.

- (24) watasi-no e-wa onnanoko-wa akai fuku-o ki-te i-masu
 I-GEN picture-TOP_{ADJ} girl-FOC_{SUBJ} red dress-ACC.OBJ wear-COMP PROG-POL
 [as for my picture, the girl is wearing a red dress]

At the next stage, the hypothesis predicts that learners can also assign prominence to nonSUBJ constituents internal to SOV such as OBJ, as in (15) above. In order to do this they need to disentangle the canonical association between the position of OBJ and its case marker, which is ACC in pragmatically unmarked sentences. In other words, OBJ marking with *-wa* as an alternative to *-o*, the ACC marker, requires full functional assignment, namely, the ability to assign a GF or DF to each constituent of the canonical word order independently of position.

In Lou's data, *-wa* marked OBJ emerges at t4, with further uses at t10 and t12. One example is shown in (25), where she first marks OBJ with *-o* (ACC) but then corrects to *-wa* (TOP). This example clearly shows that Lou knows that *-wa* may topicalise OBJ. Here we can see that, in assigning prominence to OBJ, the canonical word order is not retained.

- (25) um saakasu-no kippu-o um.. -wa ryoosin-ni morai-masi-ta
 circus-GEN ticket-ACC-TOP_{OBJ} parents-from receive-POL-PAST
 [the circus ticket, I received (it) from my parents]

In sum, Lou's data, however limited, supports the hypothesised schedule for the Prominence Hypothesis in so far as she can assign prominence to constituents by the use of *-wa* in an implicational sequence: first *-wa* marks SUBJ, the default association between a DF and a nonDF, then ADJ, a nonargument GF, and finally OBJ, a core GF other than SUBJ. Here, clearly, only the surface has been scratched, and much further work needs to be done, both in the description of Japanese native speaker's case and prominence assignment, as well as in that of learners' development.

3.2. The Lexical Mapping Hypothesis

The Lexical Mapping Hypothesis is particularly important in L2 Japanese, because many Japanese verbs and verbal constructions, including frequently used ones, require nondefault mapping of thematic roles onto GFs – represented in LFG as a-structure to f-structure mapping. Nondefault mapping may add a variety of attributions to the event such as prominence, affectedness, causality, speaker perspective, or speaker affective participation. These attributions are expressed by the speaker's choice of V – which may itself require nondefault mapping of arguments onto GFs (i.e., exceptional Vs) –, or of the V form that most closely conveys his/her intention (i.e., nonbasic V forms, such as passives, causatives and benefactives). This (lexical) mapping grammatical mechanism must not be confused with the mechanism involved in the Prominence Hypothesis, presented in the previous section, which achieves prominence by alternative linear precedence relations on the alignment of arguments – represented in LFG as c-structure to f-structure mapping.

As mentioned in chapter 1, § 2.2, this volume, default mapping, which is assumed to require least processing effort, associates agent with SUBJ and patient with OBJ. However, for discourse or pragmatic reasons, or given particular lexical choices, the speaker may highlight a thematic role other than the agent by mapping it onto the highest-ranking GF, i.e., SUBJ (Keenan & Comrie 1977). When nonagentive roles are mapped onto SUBJ, the default association of agent with SUBJ is disrupted. According to Pinker (1984) such nondefault mapping results from at least two sources: (i) intrinsically “exceptional” Vs such as *receive* in English and *morau* ‘receive’ in Japanese; and (ii) nonbasic V forms such as passives and causatives.

Thus, exceptional Vs involve language-specific V features determined by the specific lexical V selected by the speaker. Given their intrinsically language-specificity and particular behaviour, there is no systematic way of identifying them *a priori*, and learners must learn their argument specifications and mapping requirements one V at a time. Kawaguchi (2013) begins to examine the acquisition of such Vs, but further investigation is required. On the other hand, the study of the acquisition of nonbasic V forms is quite advanced for Japanese (cf., e.g., Kawaguchi 2005, 2007, 2009a, 2009b, 2010) and may open the way to their exploration in other L2s. As mentioned in chapter 1, § 4.2.2, this volume, the difference between Pinker's intrinsically exceptional Vs and nonbasic V forms is that, in the case of exceptional Vs, alternative forms, if available, are not morphologically related (e.g., *ageru/morau*, ‘give/receive’), whereas in the case of nonbasic forms, they may be morphologically derived from a more basic form which exhibits default mapping (e.g., *sikar/sikar-are*, ‘scold/be scolded’). In any case, whether or not alternative forms are lexically derived, current LFG consid-

ers them as separate lexical entries which select their own set of arguments. Thus, in (26) the passive V *sikar-are* ('be scolded') only selects the SUBJ *kodomo-ga* ('the child'), as can be seen by the grammaticality of (26a). The agent-specifying clause *sensei-ni* ('by the teacher') in (26b) turns out to complete the information and is also grammatical, but it is the speaker's choice to express it or not, because grammatically it is not essential. On the other hand, in (26c) the active V lexical entry *sikaru* ('scold') selects <SUBJ, OBJ> for its f-structure; hence *sensei-ga* ('the teacher'/agent/SUBJ) is required, as much as the other participant *kodomo-o* ('the child'/patient/OBJ) is. Of course in Japanese either may be dropped and represented by zero anaphora, but this depends on the discursive context and not on the V itself.

- (26) a. *kodomo-ga* *sikar-are-ta*
 child-NOM.SUBJ scold-PASS-PAST
 [the child was scolded]
- b. *kodomo-ga* *sensei-ni* *sikar-are-ta*
 child-NOM.SUBJ teacher-ADJ scold-PASS-PAST
 [the child was scolded by the teacher]
- c. *sensei-ga* *kodomo-o* *sikat-ta*
 teacher-NOM.SUBJ child-ACC.OBJ scold-HON-PAST
 [the teacher scolded the child]

Passives can thus be counted among lexically productive alternative V forms, that is, like other languages such as English and Italian (cf. ch. 2, § 3.2, and ch. 3, § 3.2, this volume, respectively), Japanese typically exhibits passive constructions which, among other things (such as revealing the speaker's own stance on the event or his/her affective engagement), may promote the patient role by mapping it to SUBJ (cf. the formal illustration of the nondefault mapping of a-structure onto f-structure of an English passive sentence in (26), ch. 1, this volume) and demoting the agent by mapping it onto an ADJ, as in (26b). However, the attribution of prominence is not the only function of passives. In fact there is no prominence involved at all in (26a): because there is only one argument (the theme, realised as SUBJ), there is no competition. In such cases, then, the communicative function of the passive construction is not prominence, but something else. The speaker, for instance, may want to limit the scope of the eventuality exclusively to the theme, because the agent is unknown or irrelevant. The choice of a passive construction, then, is not at all parallel to, nor to be confused with, topicalization, which realises agents as SUBJ and theme/patients as OBJ regardless of the eventual order adopted, and whose function is limited to the attribution of prominence for contrast or specification purposes.

To round up the LMH section here, I will present two other types of Japanese nondefault mapping where prominence is not directly involved, that is, benefactives and causatives. Constructions involving the AUXs ‘give’ and ‘receive’, are collectively called ‘benefactives’. Backhouse (1993: 124-125) translates three types of benefactive Vs as follows: (i) *V-te ageru* “(I/we, etc. do something for someone)”; (ii) *V-te kureru* “(someone does something as a favour to me/us)” and; (iii) *V-te morau* “(I/we, etc. receive the favour of someone doing something for us)”. Examples of these three types are shown in (27). As we can see, in benefactive constructions the lexical V precedes the benefactive AUX that contributes the benefactive role, and we notice that all three constructions here exhibit the typical SUBJ-OBL-OBJ Japanese word order for three-argument Vs.

(27) a. *give*-schema (*age*-)

Mariko-ga	kodomo-ni	hon-o	yon-de	age-ta
Mariko-NOM	child-DAT	book-ACC	read-COMP	give-PAST

[Mariko read a book to the child]

b. *give*-schema (*kure*-)

Mariko-ga	watasi-ni	hon-o	yon-de	kure-ta
Mariko-NOM	I-DAT	book-ACC	read-COMP	give-PAST

[Mariko read a book for me]

c. *receive*-schema (*mora*-)

kodomo-ga	Mariko-ni	hon-o	yon-de	morat-ta
child-NOM	Mariko-DAT	book-ACC	read-COMP	receive-PAST

[the child received the favour of Mariko reading a book]

In his LFG analysis, Ishikawa (1985) considers Japanese benefactive constructions as functionally biclausal. Later work by Matsumoto (1996) and Shibatani (1994) also support Ishikawa’s analysis. Following Ishikawa (1985: 152) and Matsumoto (1996: 48) for the representation of the lexical entry and f-structure respectively, the example of the benefactive construction in (27a) is represented in (28). As can be seen, the benefactive predicate with *ageru* (‘give’) maps in a nondefault way the benefactor and beneficiary roles onto SUBJ and OBJ respectively, and additionally the beneficial event role on XCOPM. In the f-structure SUBJ in XCOMP is linked to the same GF (i.e., SUBJ) as in the matrix clause, as indicated by the joining line.

Benefactive constructions thus display nondefault mapping because they involve a complex predicate where the f-structure of SUBJ in XCOMP can be identified by reference to the value of an argument function in a higher matrix. Without going into further detail here, also the benefactive constructions with the *morau* (‘receive’) schema present nondefault mapping, but in the opposite direction, that is, the beneficiary role is mapped onto SUBJ, the benefactive role

(28) lexical entry

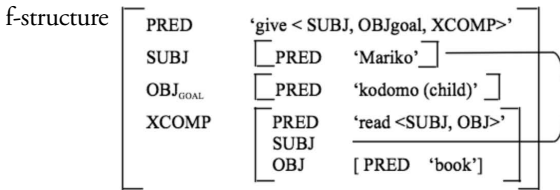
ageru 'give': V

(↑PRED) = '*ageru*

< benefactor,
(SUBJ)

beneficiary,
(OBJ_{GOAL})

beneficial event >
(XCOMP)

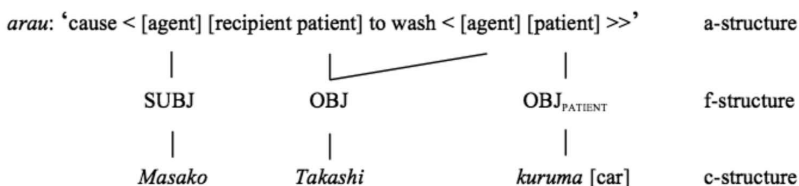


onto OBL, and so on. Given such mapping arrangements, and the complexity of structures generated by them, PT predicts that learners will produce them only once they have reached the S-procedure stage, which gives them the morphological resources to mark GFs unequivocally.

Causative constructions, similarly to benefactive ones, display a complex form-meaning relationship whereby the form expresses the meaning of 'causing X to do something or to be in some state' (Shibatani 1990: 307; cf. also Noda, Sakota, Shibuya & Kobayashi 2001). According to Matsumoto (1996), Japanese causatives involve a causer controlled sub-event where the logical SUBJ of the embedded clause is fused to the patient of the matrix clause. So, they involve non-default mapping because one participant actually receives two thematic roles. Alsina (1996: 86) assumes that "the causative verb and the base verb undergo predicate composition yielding one, single, complex, a-structure". An example of a Japanese causative construction is given in (29), more formally illustrated in (30). Here *Takashi* plays a double role in the eventuality described in the sentence: he is the patient of the causative V (*Masako made Takashi...*) and, at the same time, the agent of the lexical V *arau* (... *Takashi wash the car*).

- (29) Masako-ga Takasi-ni kuruma-o araw-ase-ta
 Masako-NOM Takashi-DAT car-ACC wash-CAUSE-PAST
 [Masako made Takashi wash the car]

(30) Mapping of a-structure onto f-structure for the transitive causative sentence Masako-ga Takashi-ni kuruma-o araw-ase-masita ('Masako made Takashi wash the car')



The syntactic hierarchy based on PT's Lexical Mapping Hypothesis for L2 Japanese is illustrated in (31), following the general proposal by Bettoni & Di Biase in chapter 1, § 4.22, (42)-(43), and similar to the language-specific hierarchies for L2 English in chapter 2, § 3.2, and for L2 Italian in chapter 3, § 3.2. In (32), Lou's data again provides the empirical evidence for the hypothesised hierarchy. Also here, like for the Prominence Hypothesis in the previous section, copula and presentational Vs are excluded from the analysis.

(31) *Developmental stages hypothesised for L2 Japanese syntax based on the Lexical Mapping Hypothesis (after Kawaguchi 2005)*

STAGE	STRUCTURE	EXAMPLE
NONDEFAULT MAPPING	causatives	<i>sensei-wa gakusei-ni takusan repooto-o kak-ase-masu</i> [teacher makes students write lots of reports]
	passives	<i>watasi-wa haha-ni sika-rare-masita</i> [I was scolded by my mother]
	benefactives	<i>tomodachi-ga kono hon-o kasite-kure-masita</i> [my friend gave me a favour of lending me this book]
DEFAULT MAPPING AND ADDITIONAL ARGUMENTS	agent/experiencer on SUBJ, patient/theme on OBJ, and other arguments on OBL	<i>watasi-wa kaban-ni keetai-o iremasu</i> [I put my mobile phone in my bag]
DEFAULT MAPPING	agent/experiencer on SUBJ and patient/theme on OBJ	<i>sensei-ga kohii-o nomimasu</i> [teacher drinks coffee]
LEMMA ACCESS	single words formulas	<i>oisii [delicious]</i> <i>arigatoo [thank you]</i>

(32) *Lou's syntactic development based on the Lexical Mapping Hypothesis (updated analysis of Kawaguchi 2005, 2009b, 2010)*

STAGE	STRUCTURE	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12
NONDEFAULT MAPPING	causative									4			
	passive								1	-1	3		
	benefactive					1-1			1	1	1		2-3
	intrinsic, nondefault						6		2-1		-1		1
DEFAULT MAPPING AND ADDITIONAL ARGUMENT	S OBL OBJ V			-1	1	1		1	4	4	1	5	3
	S OBL V		-1	4	4	6	1-1	-1		10	4-1	9	5
DEFAULT MAPPING	O V / S O V		5	4	6	7	1	4	9-1	10-1	10	18	16
	S V	1	3	2	5	2	5		1	7-1	4	4	1

- indicates wrong case marking

Lou reaches the default mapping stage with one occurrence at t1, and thereafter she shows robust numbers of positive figures with rare errors in case marking, which are indicated with a minus sign in (32). Examples of, respectively, positive and negative occurrences at t8 are shown in (33) and (34).

- (33) watasi-no okaasan-wa uhm razania-o tukuri-masi-ta
 I-GEN mother-TOP lasagna-ACC cook-POL-PAST
 [my mother cooked lasagna]

- (34) haha-wa... inu-o tabe-ta
 mother-TOP dog-ACC eat-PAST
 [literally: my mother ate the dog; intended: my mother (thinks) the dog ate (it)]

At the next stage (i.e., the default mapping and additional argument stage), the propositional content is still expressed in a pragmatically neutral, default way. This stage is reached when learners begin producing constructions consisting of default mapping together with additional arguments which are differentiated from SUBJ and OBJ, that is, when in Japanese they start using OBL arguments with case markers other than NOM and ACC. Typically, as in Lou's examples shown in (35) and (36), they are able to map the nonSUBJ argument of an intransitive V onto OBL_{LOC}, and the third argument of a transitive V onto OBL_{RECIPIENT}.

- (35) san-nen mae-ni watasi-wa nihon-ni iki-masi-ta
 three-years before-when I-TOP Japan-OBL_{LOC} go-POL-PAST
 [three years ago I went to Japan]

- (36) watasi-ni Tomu-san ah hana-o kure-masi-ta
 I-OBL_{RECIPIENT} Tom-Mr flower-ACC give-POL-PAST
 [to me Tom gave flowers]

Lou reaches this intermediate stage at t3 with four occurrences of the S OBL V structure and soon thereafter, like at the previous stage, her figures are quite consistently positive. An example of her correct case marking is shown in (37).

- (37) er fooku to naihu-o er (long pause) teeburu-ni... narabe-masi-ta
 fork and knife-ACC table-on arrange-POL-PAST
 [(I) arranged forks and knives on the table]

Finally, at the nondefault stage, learners will acquire the means to respond to pragmatic motivations that require nondefault mapping choices. At this stage, they may be able to produce benefactives, passives, and causatives. An example

of Lou's successful deployment of each nondefault mapping construction is shown in (38a-c), and one example of unsuccessful mapping of a passive construction is shown in (39).

(38) a. *benefactive construction at t9*

Jon-san-ni aisukuriimu-o kat-te ... age-masi-ta
 John-Mr-DAT ice cream-ACC buy-COMP give-POL-PAST
 [(I) bought an ice cream for John]

b. *passive construction at t9*

densya-ni not-ta toki doroboo-ni saifu-o tor-are-masi-ta
 train-on ride-PAST when robber-DAT wallet-ACC steal-PASS-POL-PAST
 [when (she) got on the train, (she) had her wallet stolen by a thief]

c. *causative construction at t10*

uh bosu-wa watasi-ni itumo kopii-o sase-masu
 boss-TOP I-DAT always photocopy-ACC CAUSE-POL
 [my boss always asks me to make photocopy]

(39) *passive construction at t9*

inu-wa .. watasi-ni... kami-rare-masi-ta
 dog-TOP I-DAT bite-PASS-POL-PAST
 [literally: a dog was bitten by me; intended: I was bitten by a dog]

In Lou's data the benefactive construction is the first to emerge at this last stage. The emergence of benefactives before passives is also confirmed by other Japanese L2 researchers (e.g., Tanaka 2001). Benefactive constructions had been taught in class just before their emergence at t5, so it is possible that teaching had an immediate if not lasting effect, because its use becomes more continuous only from t8 onwards. Lou's benefactives are in any case all *give*-schema except one, when she attempts, but fails, the *V-te morau*, *receive*-schema at t5, just after its introduction in class. Why is the *receive*-schema produced less frequently compared to the *give*-schema? Is its processing more costly? The reason may lie in the higher degree of nondefaultness with the *receive*-schema, which involves XCOMP's SUBJ linked to the OBJ source in the matrix clause, an operation that adds to the cost of online speech production. So we have here an instance of a soft barrier in the sense mentioned in chapter 1, § 4.4, this volume, and chapters 2 and 3, §§ 2.2, for L2 English and L2 Italian respectively. Within this same nondefault mapping stage, exceptional Vs emerge next in Lou's data. However, being lexically idiosyncratic and learned individually, there is no soft barrier marked in the table. On the other hand, also passives and causatives are nondefault structurally, so a soft barrier is marked in each case, respectively at t9 and t10.

In sum, naturally enough the dominant use of default mapping structures is observed throughout Lou's longitudinal study. In the first two interviews, that is in the first three months of learning Japanese for her university degree, this learner produces only default mapping structures. Then, after a further three months, besides agent/experiencer and themes mapped respectively onto SUBJ and OBJ, she is able to map additional arguments onto OBL. Finally, she attempts constructions with nondefault mapping at t5, in her second year, and uses two benefactive constructions, one of which successfully. Passive and causative constructions emerge at t9 and t10 respectively, that is, in her third year of studying Japanese for her university degree.

4. Conclusion

Japanese is one of the languages that have been most explored in PT and whose typological characteristics have contributed significantly to our understanding of the cross-linguistic plausibility of the theory, and particularly the testing of the early Topic Hypothesis and Lexical Mapping Hypothesis, as well as the current Prominence Hypothesis and the critical updating of the Lexical Mapping Hypothesis. As a matter of fact, L2 Japanese was one of two languages to provide the earliest empirical evidence for a formal typological validation of PT's morphological development – the other language being L2 Italian (cf. Di Biase & Kawaguchi 2002). Furthermore, Japanese was not only the first language to test the Lexical Mapping Hypothesis (Kawaguchi 2005), but also to do so most extensively (Kawaguchi 2010, 2013). Whereas other languages have so far provided evidence only for the development of passive structures, L2 Japanese has provided it also for benefactives and causatives, further advancing their analysis in this chapter.

This chapter presented the developmental stages in learning Japanese incorporating the latest proposals in PT, with an updated analysis of a three year longitudinal study. In particular, the Prominence Hypothesis is explored by including morphological prominence, which is characteristic of Japanese. As for the Lexical Mapping Hypothesis, it is further clarified and differentiated from the Prominence Hypothesis, and an intermediate stage is added, with relevant empirical analysis.

Broad areas within a PT framework, of course, still need attention in L2 Japanese, as well as in other L2s. These include the development of nominal morphology, interrogative sentences within the Prominence Hypothesis (a priority area to be explored in L2 Japanese), and subordination, both in morphology and syntax, as well as the syntax-pragmatics interface. In this latter area, issues such as topic continuity, nominal ellipsis and noncanonical case-marking are well worth investi-

gating from the learner's point of view. Within the Lexical Mapping Hypothesis, the acquisition of exceptional Vs requiring both nondefault mapping and non-canonical word order also needs to be addressed.

PART III

Exploring new issues

Camilla Bettoni* and Bruno Di Biase**

*University of Verona, **University of Western Sydney

Part III addresses some new theoretical and applied issues. The first five chapters explore the scope of our newly formulated Prominence Hypothesis, and deal with different languages and structures. They address the development of the case marking systems in Russian L2 and Serbian L2 (chh. 5 and 6 respectively), Differential Object Marking (DOM) in Spanish (ch. 7), constituent questions in Italian L2 (ch. 8), and V2 in German L2 (ch. 9). The last two chapters widen the scope of PT from an applied perspective and test whether PT schedules can hold in different conditions and situations, that is, the acquisition of Italian morphology by an autistic child (ch. 10), and Japanese L2 structures in a CALL mode respectively (ch. 11). A brief summary of each chapter follows.

Chapters 5 and 6 present explorations within the PT framework of the development of the case system in Russian and Serbian, two nonconfigurational, dependent-marking languages. For learners of these languages, case is a complex feature to acquire for a variety of reasons: morphologically, there are many cases, fusionally enmeshed with other nominal features such as number, gender and class. Morphosyntactically, case must be computed on most nominal elements within the NP. Then syntactically, at clause-level, case morphology itself constructs GFs independently of phrase structure. Furthermore, if relations between case and function are default and predictable most of the time, at others the same case can construct alternative GFs, and the same GF can be constructed by different cases, although with different lexical predicates. Given these complexities, the two chapters will suggest some hypotheses based on the Prominence Hypothesis, and test them out on cross-sectional data. Chapter 5 on Russian analyses the interlanguage of eight students learning their L2 in a foreign language context at the University of Verona. Chapter 6 on Serbian deals with production data of three Serbian-Australian teenage bilinguals living in Sydney, a context of contact with a (majority) language with a much-reduced expression of case, English. Results in both chapters show that there is a direct relationship between the speakers' availability of morphological case markers and their skills for deploying them to exercise alternative pragmatically driven syntactic choices. More specifically, L2 speakers seem to progress from a first match between case-form and position to full

functional assignment by case independent of position. In other words, initial learners rely on more default case markers and fixed structures for their sentences; then more advanced learners display both a fuller range of case markers, including non-default ones, and the skills for deploying them according to their discourse-pragmatic needs.

Chapter 7 addresses the development of L2 Spanish, a relatively new language to PT, with the aim of investigating the acquisition of its DOM. DOM is arguably one of the most debated topics in Spanish grammar over the last 200 years, yet little is understood in terms of its acquisition by L2 learners. There are good reasons for both these facts. Descriptive discussions must come to terms with such diverse factors as animacy of OBJ (whether animate or inanimate), specificity of OBJ (whether specific or nonspecific), form of OBJ (whether a proper N or a lexical N), and relative animacy, that is, the degree of animacy of OBJ relative to SUBJ – a discourse related ‘global’ factor. Furthermore, descriptions often fail to take into account sociolinguistic variation of the use of DOM over many countries where Spanish is spoken in a variety of settings (monolingual, bilingual, heritage, etc.). In the acquisition literature, DOM is often treated as a purely structural phenomenon which is supposedly easy to learn. Chapter 7 instead places DOM high up among the last stages of PT, and explains why that is so.

Constituent questions are dealt with in chapter 8. They are extremely interesting for PT in so far as they are sentences marked both pragmatically and grammatically. Pragmatically, they satisfy an important communicative need, as speakers use them to request new information. Hence constituent questions always have an element ‘in focus’ (Lambrecht 1994; Mycock 2007), which is the interrogative phrase. Focality is not a prerogative of questions, but in this type of sentence the focus is obligatory, and responds to specific linguistic constraints. Hence, constituent questions are also linguistically marked. Such constraints vary cross-linguistically, and can be structurally complex to encode. Furthermore, in comparison with declarative sentences, constituent questions occur less frequently in spontaneous conversation. For all these reasons, it is not surprising that they are difficult to acquire by L2 learners. Chapter 8, then, describes how content questions are realized syntactically in Italian by using the LFG framework, then discusses PT’s hypotheses for their developmental hierarchy based on the Prominence Hypothesis, and finally tests these hypotheses on empirical cross-sectional data from learners with a variety of L1 backgrounds.

Chapter 9 deals with German declarative sentences and constituent questions in relation to its V2 rule, which is seen from the new PT perspective of the Prominence Hypothesis. This chapter analyses learners’ development beyond canonical word order, and compares progress in topicalised declarative sentences and in constituent questions. Results show that V2 – that is, noncanonical word

order – emerges in questions before it does in declaratives, thus suggesting that question FOC is a more powerful trigger than TOP in learners' progress beyond canonical order.

Chapter 10 deals with an autistic child acquiring Italian L2. Among the defining features of autism are delays and deficits in language and communication. However, the exact nature of these problems is unclear because language outcomes vary greatly. For example, some children never acquire speech, others acquire only limited speech and yet others, despite early delays, acquire language within the normal range. Little is known about grammatical development in high functioning children, yet understanding how language develops in the early stages in this population may provide valuable insight into how underlying processing difficulties contribute to their speech delays. This chapter presents a case study designed to assess whether a 6-year old child with high functioning autism learns an L2 by the same developmental path as typically developing children. Results confirm that it is the same, with some indication that progress is not slower. Rather, each stage seems to develop at a rapid pace.

Chapter 11 addresses pedagogic and research methodological issues by connecting CALL and PT with a focus on Japanese L2. The CALL activity examined is text-messaging exchange between learners of Japanese L2 in Australia and learners of English L2 in Japan. Participants text-chatted on various topics as an out-of-class activity three times over a period of two months. Japanese L2 production during these sessions is analysed in terms of lexicon, morphology and syntax. Furthermore, language development is examined in order to check whether the trajectory of morphosyntactic structures defined by PT for oral production is confirmed also for the written production during text messaging. Results confirm Pienemann's (1998) Steadiness Hypothesis, but they also suggest that there are vast individual differences in students' language productions and learning outcomes as measured by PT stages. This points to the need of not only monitoring learners closely to promote overall linguistic development, but also using a reliable developmental measure such as PT's schedules.

Acquiring case marking in Russian as a second language: an exploratory study on subject and object

Daniele Artoni and Marco Magnani

University of Verona

1. Introduction

This chapter deals with the development of case in Russian as a second language. Case is an important morphological device for marking grammatical relations among constituents, and as such is of great interest for testing the interface between morphological and syntactic development proposed by Bettoni & Di Biase in the introductory chapter of this volume (ch. 1, § 4.3). Yet so far, few studies have specifically dealt with the L2 acquisition of case. Among the exceptions, some are notable. Baten's (2011, 2013) study on the acquisition of German case, and the study on Serbian case in chapter 6 of this volume are within the PT framework. Although they are different in design – the former being longitudinal on a variety of Dutch L1 learners of German L2, and the latter cross-sectional on learners of Serbian as a heritage language in Australia – they both show that learners begin to mark case first positionally in the clause (i.e., they assign NOM to preverbal SUBJ, and ACC to postverbal OBJ), and then functionally (i.e., regardless of position in c-structure). Without the PT framework, studies on the L2 acquisition of case are also scarce generally, and even more so with respect to Russian. One exception is Kempe & MacWhinney (1998), who investigate the acquisition of case marking in Russian in comparison to German. Their findings are interesting in so far as they show that, although the Russian case-marking system is more complex than the German one, it appears to be learned faster because its inflections are more reliable for sentence interpretation. However, their study deals with comprehension, and production remains unexplored.

In line with the PT framework proposed in this volume, in our chapter we explore the development of Russian case and hypothesise a developmental schedule interfacing morphology with syntax. More specifically, we focus on how case morphology is used to mark GFs, with particular attention to the two core GFs, namely, SUBJ and OBJ. Our hypotheses are then tested on cross-sectional data col-

lected among 8 learners of Russian L2. Results will show that learners progress from a first match between GFs and their default case markers in a fixed canonical word order frame to expressing GFs also with nondefault case markers as lexically required by V, and finally to marking full functional assignment by case independent of position. This allows learners to deploy more flexible word orders to express their discourse-pragmatic needs unambiguously.

2. Case in Russian

From a typological point of view, Russian is a language with a low degree of configurationality and a rich case morphology that set it among the dependent-marking languages (for an overview of this language, cf. Comrie 2011). A case system is a prominent characteristic of dependent-marking languages, traditionally defined, in a general way, as a system marking dependent nominals to the type of relation they bear to their heads in a phrase (Blake 1994). Case, then, is not a universal feature, as GFs can be identified by three different means: (a) case marking, which is the main means used by dependent-marking languages such as Russian, Japanese (cf. ch. 4, this volume), and Warlpiri (cf. (24), § 2.2, ch. 1); (b) agreement, which is very productive in Romance languages like Italian (cf. ch. 3) and Spanish (cf. ch. 7), where SUBJ and V must agree in person, number, and sometimes also gender; and (c) word order or position in phrase structure, used in configurational languages like English (cf. ch. 2). Natural languages can then obviously exploit more than one means to identify GFs. Russian and Latin, for instance, present both a rich case morphology and SUBJ-V agreement. This allows speakers to resort to different word orders and organise sentences according to the pragmatic requirements of the TOP-FOC structure of their sentences (cf. §§ 3.1, ch. 3 on Italian, and ch. 4 on Japanese).

Within the LFG framework, King (1995) discusses four types of case assignment in Russian: semantic, configurational, functional, and lexical. Semantic case assignment, as its label suggests, occurs when a particular case is associated with a particular semantic meaning at a-structure. Semantic cases are common across languages, but according to King (1995) the only candidate in Russian is the instrumental case for <instrument>, as shown in (1).

- (1) ja napisala pis'mo karandašom
 I-NOM wrote letter-ACC pencil-INST
 [I wrote a letter with a pencil]

Configurational case assignment occurs when a specific case is assigned to a N appearing in a certain position in phrase structure. In King's view, this occurs in

Russian when genitive in N is daughter of NP, as exemplified and formalised in (2). Notice that, unlike with semantic case, genitive is assigned only by position in c-structure because the genitive sister of N can mark different semantic roles, such as possessor, as well as agent as in (2).

- (2) a. otvet učenika
 answer-NOM pupil-GEN
 [(the) answer by (the) pupil]
- b. NP → N (NP)
 ((↓CASE) = GEN)

Crucially, to build up the sentence, case assignment can mark GFs. In Russian three GFs require their own default case, namely, nominative (NOM) for SUBJ, accusative (ACC) for OBJ, and dative (DAT) for OBL_{GOAL}, as exemplified and formalised in (3).

- (3) a. mal'čik dal Inne knigu
 boy-NOM gave Inna-DAT book-ACC
 [the boy gave Inna a book]
- b. (↑SUBJ CASE) = NOM
 (↑OBJ CASE) = ACC
 (↑OBL_{GOAL} CASE) = DAT

Finally, lexical case assignment occurs when case is governed by a particular verb or preposition, as exemplified and formalised for verbs in (4) and for prepositions in (5) respectively. Note that the rule in (4) is not contradicted by those in (3) because specific lexical requirements can override defaultness.

- (4) a. upravljaet biznesom
 manages business-INST
- b. *upravljat'* 'manage' V <SUBJ, OBJ>
 (↑OBJ CASE) = INST
- (5) a. u okon
 by windows-GEN
- b. *u* 'at/near' PREP <OBJ>
 (↑OBJ CASE) = GEN

In this study we will deal with GF case assignment and lexical case assignment by V.

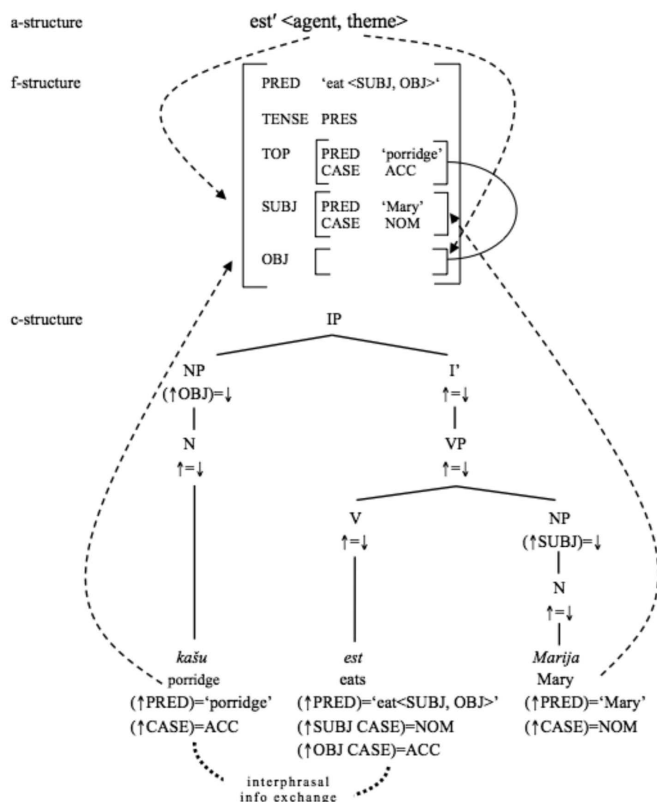
Russian canonical word order is SVO,¹ so SUBJ marked as NOM is the default TOP in preverbal position, and OBJ marked as ACC is the default FOC in postverbal position. However, because of Russian nonconfigurationality, SVO occurs only 47% of the time in native oral production (Timberlake 2004). This means that case-marked GFs are not positionally predictable. For discourse and pragmatic reasons, constituents can occur in different positions, allowing for all the six possible combinations of the three core elements, as shown in (6) – even though it is important to note that word orders in (6d-f) rely heavily on prosodic features and, being highly marked, are rarely used by Russian L1 speakers (Kallestinova 2007). In the examples below these six combinations are identified thanks to the feminine nominative *–a* marking SUBJ, and the feminine accusative *–u* marking OBJ.

- (6) a. Marija est kašu
 Marija-NOM.SUBJ eats-V porridge-ACC.OBJ
- b. Marija kašu est
 Marija-NOM.SUBJ porridge ACC.OBJ eats
- c. kašu est Marija
 porridge-ACC.OBJ eats-V Marija-NOM.SUBJ]
- d. kašu Marija est
 porridge-ACC.OBJ Marija-NOM.SUBJ eats-V
- e. est Marija kašu
 eats-V Marija-NOM.SUBJ porridge-ACC.OBJ
- f. est kašu Marija
 eats-V porridge-ACC.OBJ Marija-NOM.SUBJ

What unites the sentences in (6) is that they all share the same *f*-structure. On the other hand, the most relevant difference in *c*-structure is that OBJ occurs within VP when it is postverbal, whereas it is set outside VP when it is displaced to TOP position (King 1995: 206), as shown in (7). This implies that when OBJ is preverbal, the value of the OBJ CASE feature of V must be checked interphrasally with the case value of NP_{OBJ}.

1 For an alternative view on Russian word order, cf. King (1995), who suggests VSO as the unmarked, pragmatically neuter word order.

(7) Correspondences between a-structure, f-structure and c-structure for the sentence *kašu est Marija* ('the porridge Marija eats (it)')



Readers will appreciate that for learners of Russian L2 case is a complex feature to acquire. Yet the complexity of its functional deployment is well matched by the complexity of its formal characteristics. There are six cases in Russian: nominative, genitive, dative, accusative, instrumental, prepositional (locative and ablative), which are fusionally enmeshed with other nominal features such as number (singular or plural), gender (masculine, feminine or neuter), animacy, and class. In (8) and (9) we show the full case-marking paradigm for Ns and pronouns respectively. As we can see, the many-to-many relations between cases and markers are noteworthy. In particular, with regard to SUBJ and OBJ, the two cases of interest in this chapter, neuter and masculine inanimate Ns share the same marking strategies for NOM and ACC. For this reason, we will consider only feminine and masculine animate Ns for matching SUBJ to NOM, and OBJ to ACC. On the other hand, with regard to pronouns, no ambiguity between markers is found in the NOM and ACC declension.

(8) *Russian case-marking paradigm: nouns (after Kempe & MacWhinney 1998)*

	S I N G U L A R				
	MASCULINE		NEUTER	FEMININE	
	ANIMATE	INANIMATE		1 ST CLASS	2 ND CLASS
NOM	-ø	-ø	-o/-e	-a/-ja	-'ø
GEN	-a/-ja	-a/-ja	-a/-ja	-y/-i	-i
DAT	-u/-ju	-u/-ju	-u/-ju	-e	-i
ACC	-a/-ja	-ø	-o/-e	-u/-ju	-'ø
INST	-om/-em	-om/-em	-om/-em	-oj/-ej	-'ju
PREP	-e	-e	-e	-e	-i

	P L U R A L				
	MASCULINE		NEUTER	FEMININE	
	ANIMATE	INANIMATE		ANIMATE	INANIMATE
NOM	-y/-i	-y/-i	-a/-ja	-y/-i	-y/-i
GEN	-ov/-ev/-ej	-ov/-ev/-ej	-ø /-ej	-ø/-ej	-ø/-ej
DAT	-am/-jam	-am/-jam	-am/-jam	-am/-jam	-am/-jam
ACC	-ov/-ev/-ej	-y/-i	-a/-ja	-ø/-ej	-y/-i
INST	-ami/-jami	-ami/-jami	-ami/-jami	-ami/-jami	-ami/-jami
PREP	-ach/-jach	-ach/-jach	-ach/-jach	-ach/-jach	-ach/-jach

When -o and -a are not stressed, they are both pronounced /ə/

(9) *Russian case-marking paradigm: pronouns (after Kempe & MacWhinney 1998)*

	S I N G U L A R						P L U R A L		
	PERS	1 ST	2 ND	3 RD			1 ST	2 ND	3 RD
				MASC.	NEU.	FEM.			
NOM		ja	ty	on	ono	ona	my	vy	oni
GEN		menja	tebja	ego	ego	eë	nas	vas	ich
DAT		mne	tebe	emu	emu	ej	nam	vam	im
ACC		menja	tebja	ego	ego	eë	nas	vas	ich
INST		mnoj	toboj	im	im	ej	nami	vami	imi
PREP		mne	tebe	nëm	nëm	nej	nas	vas	nich

Pronouns beginning with a vowel may be preceded by an epenthetic -n

3. Developmental Hypotheses

In (10) we show our developmental hypotheses for Russian case relative to the GFs SUBJ and OBJ. As we have noted in the introduction to this chapter, case is particularly interesting for PT because the two schedules for morphological and syntactic development interface in a crucial way. So, contrary to other chapters in this volume (e.g., chh. 1-4), we include both morphological and syntactic development

in a single table. We should also mention that our Prominence Hypothesis deals with declarative sentences, leaving interrogatives to future work.

(10) *Developmental stages hypothesised for Russian case*

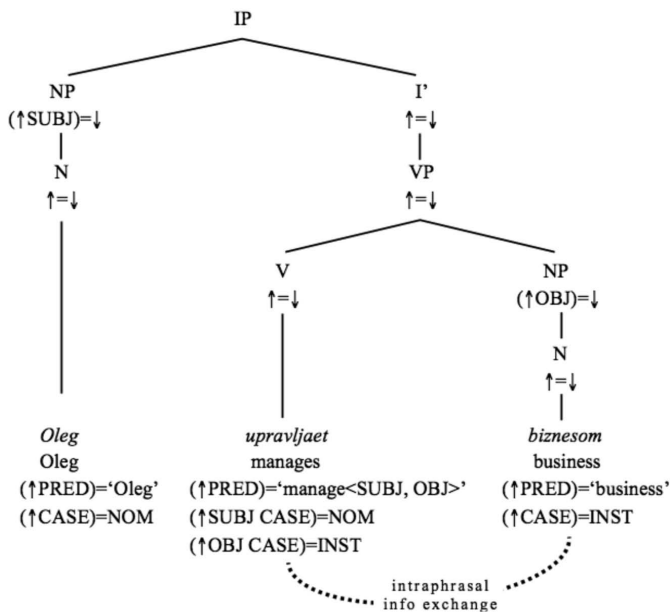
	STAGE	STRUCTURE	MORPHO-SYNTACTIC OUTCOME	EXAMPLE
	SYNTAX	NONCANONICAL WORD ORDER	OVS, OSV, etc.	
				<i>knigu čitaet mama</i> [book-ACC reads mum-NOM]
	MORPHOLOGY	SENTENCE PROCEDURE	TOP _{OBJ} & V unification SUBJ & V agreement	
	SYNTAX	XP _{DF} CANONICAL WORD ORDER	TOP _{ADJ} SVO	<i>sejčas Oleg smotrit televizor</i> [now Oleg-NOM watches television-ACC]
	MORPHOLOGY	PHRASAL PROCEDURE	V & OBJ unification	<i>upravljaet biznesom</i> [manages business-INST]
	SYNTAX	CANONICAL WORD ORDER	SVO	
				<i>devočka est' kašu</i> [girl-NOM eat porridge- ACC]
	MORPHOLOGY	CATEGORY PROCEDURE	case marking on N: NOM vs non-NOM	
	LEMMA ACCESS	single words formulas		<i>net urok</i> [no lesson] <i>menja zovut Sergej</i> [my name is Sergej]

After the single-word and formulaic stage, as soon as the category procedure becomes operative for morphology, learners are able to distinguish categorially between Ns and Vs. Formal marking of Ns, then, begins to emerge. With regard to the case feature specifically, at this stage learners begin to distinguish between the NOM form and a general non-NOM form bearing any inflectional ending other than NOM (e.g., *kaša* 'porridge-NOM' vs. *kašu/kašel/kaši* 'porridge-non-NOM'). This initial opposition between NOM and non-NOM (cf. Jakobson 1971), although minimal and restricted within the word, is sufficient to distinguish formally between SUBJ and non-SUBJ GFs in syntax. However, because at this stage GFs are identified positionally rather than functionally (cf. ch. 1, § 4.2.1), learners will deploy this minimal form variation onto a minimally specified SVO sequence – that is, they will produce canonical strings with a preverbal N_{NOM} and a postverbal N_{non-NOM}, as shown in (11).

- (11) ochotniki ubit volke
 hunters-NOM kills wolf-non-NOM
 [the hunters kill the wolf]

At the next stage, with regard to morphological development, learners are able to exchange information within a phrase. In Russian this phrasal procedure stage involves a variety of structures, such as agreement between noun and adjective within NP and government by the preposition within PP. The relevant structure for our study is V-OBJ unification when learners are able to exchange information within the VP between the value of the case feature of NP_{OBJ} and the value of the feature OBJ CASE required by V. If OBJ is marked by its default ACC case in its default postverbal position, the evidence of intraphrasal exchange of information remains equivocal. We prefer to consider unequivocal evidence of progress to this stage OBJs marked with cases other than ACC (e.g., INST)², as exemplified above in (4), § 2, and further formalised in a full sentence in (12). Needless to say, this does not imply that all nondefault cases required by V will be learned soon. Being required lexically, they will have to be learned individually V by V.

- (12) *Annotated c-structure for the sentence Oleg upravljaet biznesom ('Oleg manages business')*



2 Also OBJ_{INST} could randomly appear as OBJ_{non-NOM} at the category procedure stage. What provides unequivocal evidence for online information exchange within the VP is a variety of cases with different lexical Vs.

With regard to syntax, we hypothesise that at this stage learners will be able to place an element other than SUBJ (typically ADJ) in the prominent first position as in (13). This addition will bring about a differentiation between SUBJ_{NOM} and the topical first constituent in the clause.

- (13) na kartinke devuška est jabloko
 in picture girl-NOM eats apple-ACC
 [in the picture the girl eats an apple]

At the last stage of their morpho-syntactic development, learners will be able to assign GFs irrespectively of position. As we have seen, this requires two morphological resources in Russian: a head-marking strategy, namely SUBJ-V agreement for the identification of SUBJ; and a dependent-marking strategy, namely case-marking for identifying the three main argument functions – that is, SUBJ, OBJ and OBL_{GOAL}. With regard to the former strategy, with the activation of the S-procedure, learners can now produce the agreement between the SUBJ features (number and gender) and the predicate, as shown in (14). Although SUBJ-V agreement is not the focus of this study, in (14) we show the unification of SUBJ with a nonverbal predicate rather than with a lexical V because in Russian, as well as in other null-SUBJ languages,³ both SUBJ and V need to retrieve their features from the conceptual structure independently (cf. ch. 3, § 2.1 on Italian L2 for a discussion on this point).

- (14) pogoda byla chorošaja
 weather-SG.FEM was-SG.FEM good-SG.FEM
 [the weather was good]

Thanks again to the activation of the S-procedure, information exchange between V and its complements can now occur across phrases. Hence learners will be able to case-mark the topicalised OBJ with ACC as a result of the exchange of information between VP and the external NP, and produce sentences like (15).

3 Russian, however, is not universally recognised as a null-SUBJ language. For instance, Avrutin & Rohrbacher (1997) argue that Russian is not a null-SUBJ language, and ascribe the rare instances of null SUBJ to contextually licensed ellipsis. On the other hand, Müller (1988) and Pearlmutter & Moore (2002) support the null-SUBJ thesis, and justify the contextual limits on the basis of discourse conditions that “make it much less common than pro-drop in Italian or Spanish”.

- (15) knigu čitaet mama
 book-ACC reads mum-NOM
 [the book mum reads (it)]

At this last stage, then, the interplay between morphology and syntax in the development of case is clear. On the one hand, morphology feeds syntax in the sense that only when the S-procedure is firmly in place can learners case-mark constituents unambiguously regardless of word order constraints. On the other hand, along the path for morphological development, when learners are able to free up the rigidity of the canonical word order frame – crucially, by choosing to topicalise the core function OBJ – they provide convincing evidence that case is assigned via interphrasal information exchange.

4. The study

The evidence we bring to test our PT-based hypotheses discussed in § 3 comes from a cross-sectional study of 8 learners of L2 Russian (2 males and 6 females), all adults at different proficiency levels. As the table in (16) shows, they are all instructed learners, who have studied Russian at university for one to five years. Two of them (AB and LI) have also acquired and practised the language in a Russian speaking environment. Their L2 proficiency levels vary from A2 to C1 as measured on the general CEFR (2001) scales. Their L1 backgrounds are also varied, including Italian for the majority of them, Serbian (JO) and Georgian (LI). It is however important to note that, although some of the variables in (16) – particularly the L1 background – can contribute in speeding up progress along the developmental path, they will not affect the sequence in which stages are reached (cf. the Developmentally Moderated Transfer Hypothesis in Pienemann, Di Biase, Kawaguchi & Håkansson 2005).

(16) *The learners*

LEARNER	AGE	GENDER	L1	L2 INSTRUCTION	L2 IMMERSION	L2 PROFICIENCY
AB	24	M	Italian	5 years	5 months	C1
AL	25	F	Italian	3 years		B2
CA	26	F	Italian	3 years		B1
EL	20	F	Italian	2 years		B1
JO	22	F	Serbian	2 years		B1
LI	25	F	Georgian	4 years	24 months	C1
MA	20	F	Italian	1 years		A2
MT	22	M	Italian	3 years		B2

We used five tasks to elicit the structures listed in (10). The first task, *Znakomstvo* ('Introduction'), is a conversation aimed to get to know and relax the learners, and targeting general structures. The next two tasks, *Krasnaja Šapočka* ('Red Riding Hood') and *Najdite različija* ('Spot the differences'), are respectively a story telling and spot-the-difference task targeting canonical word order and ADJ topicalisation. In the fourth task, *Eščë Krasnaja Šapočka* ('Red Riding Hood Again'), two cards showing a person and an object were introduced together with a V in its infinitive form, and the learner was asked to create a sentence with the three given items. This task targets the use of OBJ with cases other than ACC which are lexically required by V, such as *upravljat'* ('manage') and *zanimat'sja* ('do/practice'), both requiring OBJ to be marked by INST. The last task, *Večerinka* ('The party'), targets OBJ topicalisation, and is an adaptation to Russian L2 of the task used first by Di Biase & Kawaguchi (2002) and then by Bettoni & Di Biase (2011) for eliciting Italian L2 structures (cf. § 2.2, ch. 3, this volume). Di Biase's 'animal dinner' task becomes a 'party' to which various participants must contribute something. When two picture cards appear on the computer screen, learners are encouraged to tell the researcher who is bringing what by starting with the card on the left. Since this card shows sometimes the person/agent, other times the object/theme, SVO sentences should alternate with OVS ones. The communicative context of all the five tasks and the use of several distractors scattered among the targeted structures aim to enhance learners' online production, and exclude a focus on declarative knowledge.

Our analysis comprises a total of 476 case-marked structures included in 333 main declarative sentences with lexical Vs – thus excluding copular and presentative sentences, which are 'nonverbal predicates' (Kroeger 2005: ch. 10). In (17) we illustrate the distributional analysis of the case markers on SUBJ and OBJ among learners, structures and stages. The table is organised as set out in this volume (cf. (48a-c), § 5, ch. 1), but four further points must be clarified. First, most structures in the first column are repeated twice because when both GFs are case-marked in a sentence, the figures for case refer only to the GF in bold, and the brackets around the other GF mean either that its case is not considered in that row or that the GF is absent from the string (there being a null SUBJ or an intransitive V). This then explains why in the table the total number of case-marked GFs (476) is higher than the number of analysed clauses (333). Secondly, given the relatively high degree of case syncretism in the Russian case system, in our analysis we consider only unambiguously case-marked Ns. So, with regard to the marking of NOM and ACC, as we have anticipated in § 2, for evidence of acquisition of ACC we only look at case marking on feminine nouns (e.g., *ručka* 'pen-NOM' vs. *ručku* 'pen-ACC') and masculine animate nouns (e.g., *student* 'student-NOM' vs. *studenta* 'student-ACC'), and thus exclude masculine inanimate and neuter nouns, where NOM and ACC are equal-

ly marked. With regard to the marking of INST for lexical case assignment, there being no ambiguity with other case marks, we consider all nouns, provided that the structure is produced online in a nonformulaic way. Thirdly, as evidence of acquisition of V-OBJ unification through lexical case assignment, we do not consider the occurrences of OBJs marked by the default ACC instead of the non-default INST, because the use of the default ACC here may be due to the lack of annotation of the OBJ CASE feature in the verbal lexicon. Thus, among the figures for OBJ_{INST}, the minus always indicates that OBJ is marked by NOM – an unequivocal example of lack of information exchange with V. Fourthly, following Pienemann (1998: 159), in this analysis we factor in the case feature only, and consequently factor out number and gender. This means that, for example, in the case of OBJ marked by INST, the use of the INST masculine ending *-om* in **babušk^{om}* ('grandmother-INST') instead of the correct feminine one *-oj* is considered as safe evidence of acquisition of the structure.

(17) *Cross-sectional study of the development of Russian case*

STRUCTURES	AL	MA	JO	EL	CA	LI	AB	MT
OBJ_{ACC} V (SUBJ)	-4	-4	-3	-4	-4	-5	+4	+4
(OBJ) V SUBJ_{NOM}	(+3)	(+1 -3)	(+1 -2)	(+4)	(+4)	(+4)	(+4)	(+4)
ADJ (SUBJ) V OBJ_{ACC}					-1	+4	+4	+2 -1
ADJ SUBJ_{NOM} V (OBJ)					+3	+6	+9	+6
(SUBJ) V OBJ_{INST}	-1	-1		+1	+1 -1	+2	+2	+2
(SUBJ) V OBJ_{non-NOM}	+8 -9	+7 -2	+15	+6 -8	+3 -2	+14 -2	+9 -8	+12 -3
SUBJ_{NOM} V (OBJ)	+26	+32	+31	+40	+25	+32	+45	+28

Numbers in brackets are irrelevant for determining learners' progress; cf. the analysis

All our learners can produce sentences of preverbal Ns marked as NOM and postverbal Ns marked as non-NOM, as in (18). Their numbers are convincing, and we can thus safely say that they have all reached the category procedure stage for morphology and the canonical word order stage for syntax.

- (18) AL videla volke
 saw wolf-non-NOM

However, a finer analysis reveals interesting differences in how OBJ is marked. On the one hand, the overall average for the use of the incorrect NOM marker is 31.48%. All learners but one (JO) use it, even the more proficient ones, as shown by the minus figures for AB (9 out of 17) and exemplified in (19). On the other

hand, when OBJ is marked by a case other than NOM, the overall average for the correct ACC *-u/-ju* marker is 93.24%, as exemplified in (20). All learners but one (AL) use ACC all the time, whereas AL, the least proficient, alternates between the ACC marker and the PREP case *-e* marker, as in (18) above.

- (19) AB ochotniki našli volk
 hunters-NOM found-PL wolf-NOM
 [the hunters found the wolf]

- (20) LI oni uvideli volka
 they-NOM saw-PL wolf-ACC
 [they saw the wolf]

Four of our learners, CA, LI, AB and MT, show evidence of having reached the next stage, which for morphology involves VP unification between V and OBJ. As we have mentioned in § 3, uncontroversial evidence of this unification can be captured when V requires OBJ marked by a case other than ACC, such as INST in (21). However, only one such instance by EL does not satisfy our emergence criterion for morphology (cf. § 5, ch. 1, this volume).

- (21) MT ona z анимаetsja muzykoj
 she-NOM does music-INST
 [she practices music]

These four learners also use structures that topicalise ADJ, as in (22). Notice that evidence of progress at this stage is determined by the syntactic position of SUBJ_{NOM}, which is no longer the default (initial) one, rather than by its case marker, which is the default NOM.

- (22) LI potom ona posmotrela babušku
 later she-NOM saw grandmother-ACC
 [then she saw her grandmother]

Two learners, AB and MT, have also reached the last stage of development in so far as they are able to use case to mark GFs irrespectively of their position in the clause. Hence, they can produce sentences like (23), where OBJ marked as ACC is in pre-verbal position.

- (23) MT vilku prinesla balerina
 fork-ACC brought dancer-NOM
 [the fork, the dancer brought it]

As we have remarked earlier, OBJ topicalisation in these constructions was deliberately prompted by our fifth task. What happens, then, when learners who have not yet acquired the morphological resources of the last stage are asked to perform this task? The most common solution is to use NOM on both Ns in the OVS structure, as shown in (24). Although here postverbal NOM is correct, we cannot take it as evidence of progress to this noncanonical word order stage because, as we have remarked above, NOM is the default case – hence the brackets in the table.

- (24) AL vilka prinës balerina
 fork-NOM brought dancer-NOM
 [? the fork brought the dancer/the dancer brought the fork]

One could argue that lack of case marking in (24) may be motivated by the fact that AL's L1 is Italian, a language which never marks Ns by case. However, we can safely say that this is not an issue of L1 transfer, as ACC is never marked on topicalised OBJs even by JO and LI, whose L1s are Serbian and Georgian respectively – two heavily morphologised languages with a rich case system. In this regard, we note that in two out of three of these topicalised declaratives, the Serbian learner JO marks preverbal OBJ by the default NOM, and postverbal SUBJ by ACC, as shown in (23).

- (23) JO vilka prinës balerinu
 fork-NOM brought dancer-ACC
 [the fork brought the dancer]

This is a common solution for learners who are at the earlier developmental stage. MA, for example, marks 3 out of 4 postverbal SUBJs as ACC. So it would seem that for 6 of our learners syntactic position still leads the way for marking case, even in the context of an L1 background which is typologically very similar to Russian.

5. Discussion and conclusion

In this chapter we hypothesised a developmental hierarchy for case in Russian L2 considering the interface between PT's schedules for morphology and syntax. We have seen how case presents an intriguing challenge for testing such an interface, and how Russian offers an ideal testing ground for several reasons – chiefly among them its rich morphological system and its high nonconfigurationality, allowing for permutations of the core elements in the clause.

Results of our cross-sectional study confirm our PT-based hypotheses, in so far as on both the morphological and the syntactic schedules there is no learner who produces structures at a higher stage without producing also at least some at the previous stage. Results also support the interface between morphological and syntactic development suggested by Bettoni & Di Biase in chapter 1, § 4.3. In particular, the interface is clear at the lowest and the highest grammatical stages. That is, first, when the category procedure becomes operative in morphology, learners begin to distinguish between NOM and a general non-NOM form, but can only deploy this minimal form variation within the fixed frame of the SVO canonical word order; and then last, only when learners can activate the morphological resources of the S-procedure can they free up the rigidity of canonical word order constraints, and assign case to GFs irrespectively of position. On the other hand, in the intermediate stages, that is, when the phrasal procedure stage is supposed to interface with the XP_{DF} canonical word order stage, the interface between the learners' ability to produce V-OBJ unification in morphology and to topicalise an ADJ in first position seems less clear, as also hypothesised by Bettoni and Di Biase in chapter 1, § 4.3.

A closer look at learners' production of pragmatically-driven OVS sentences can further contribute in explaining the path from case assignment based on position towards case assignment irrespectively of position. Three main outcomes are used by learners in our data: (i) both NPs are marked by the default NOM; (ii) both NPs are marked positionally, that is, TOP_{OBJ} by NOM and postverbal SUBJ by ACC; and (iii) like in target Russian, both NPs are case-marked functionally regardless of position, that is, TOP_{OBJ} by ACC and postverbal SUBJ by NOM. Such outcomes are not randomly distributed across learners, and may be interpreted as subsequent stages along the developmental path. More specifically, overextension of NOM on both NPs tends to occur both at the beginning of the interlanguage path (AL) as in (24), and at intermediate stages (EL, CA and LI) as in (25).

- (24) AL gitara prinēs balerina
 guitar-SG.FEM brought-SG.MASC dancer-NOM.SG.FEM
 [? the guitar brought the dancer / the dancer brought the guitar]

- (25) LI gruša prinesla prepodavatel'nica
 pear-NOM.SG.FEM brought-SG.FEM teacher-NOM.SG.FEM
 [? the pear brought the teacher / the teacher brought the pear]

In (25), however, there is more accuracy in V morphology than in (24), in so far as V is correctly inflected as SG and FEM – even though it is impossible to tell whether it agrees with preverbal N or with postverbal N. Moreover, although

both beginner and intermediate learners overextend NOM in some sentences with canonical word order as well, we have seen that intermediate learners are more accurate in using ACC for OBJ than AL, who uses a variety of non-NOM markers for OBJ. Thus, whereas AL's overextension of NOM can be seen as a general default solution, for intermediate learners it seems to be an indicator for their inability to use NOM and ACC in ways other than positionally. A longitudinal study would confirm this whole sequence for case, as well as for SUBJ-V agreement in all learners.

Our study also shows that typological similarities between the L1 and the target L2 do not allow learners to skip stages along the developmental sequence. In fact, our Serbian learner (JO) also marks constituents by case according to their position and is thus at the earlier developmental stage. On the other hand, our analysis suggests that such similarities can play an important role in increasing accuracy within a stage. In fact, JO is the only one who marks postverbal OBJs always accurately with ACC rather than with a non-NOM case.

Finally, using King's (1995) terms for case assignment, we have resorted to evidence of lexical case assignment for proving V-OBJ unification at the phrasal procedure stage, and to grammatical case assignment for proving TOP_{OBJ}-V unification at the S-procedure stage. This of course does not entail that lexical case markers are acquired before structural ones generally. This study makes no predictions as to the point of emergence of the different lexical case markers – a thorny issue which needs to be worked out at the interface with the Lexical Mapping Hypothesis. It only uses them when they are found in the data as evidence for the activation of the interphrasal procedure in order to rule out the possibility that default ACC is assigned simply because of the N position.

In conclusion, this study has shown that marking SUBJ and OBJ by case in Russian is no mean feat for L2 learners. Although minimal variation between NOM and non-NOM is sufficient for distinguishing between SUBJ and OBJ in the fixed canonical word order frame, trouble begins when V lexically requires a nondefault match between GF and case, and when discourse-pragmatic requirements place GFs other than SUBJ in the prominent first position. Indeed, only intermediate learners will start marking OBJ by the nondefault INST case, and only the most proficient ones can mark the GFs morphologically even when sentences display noncanonical word order. PT can explain why this is so by tracing the learner's developmental path from case marking based on position towards case assignment independently of position. On the other hand, our study only scratches the surface of the acquisition of case in a highly nonconfigurational language such as Russian. Further investigation is needed in several directions, as well as more substantial evidence on more diverse structures in a wider corpus. Future work, for example, should include the third case-

marked argument GF, namely the dative OBL_{GOAL} ; interrogative sentences as well as the interface between PT's Prominence Hypothesis and Lexical Mapping Hypothesis with regard to syntax; and King's (1995) semantic and configurational case assignments with regard to morphology. These are unexplored areas in Russian L2.

The authors wish to thank Camilla Bettoni, supervisor of their doctoral work, and Bruno Di Biase, her co-editor of this volume, for their help in discussing and organising the issues dealt with in this chapter.

6

The development of case: a study of Serbian in contact with Australian English

Bruno Di Biase*, Camilla Bettoni** and Lucija Medojević*

*University of Western Sydney, **University of Verona

1. Introduction

Within the PT framework, this chapter represents an exploration in three new directions. First, we look at the development of a case system in learners. Despite PT's fairly extensive empirical application to a variety of languages and acquisitional circumstances and populations (cf. § 1, ch. 1, this volume), surprisingly little attention has been given to the development of case systems, except for Baten's work on German L2 (2011, 2013) and Artoni & Magnani's on Russian L2 (2013) and in chapter 5, this volume. The system we will be looking at is that of Serbian, like Russian a heavily morphologised case-marking language (for an overview of Serbian from a typological point of view, cf. Corbett & Browne (2011). After a brief account of the case system in Serbian, we will present our PT-based hypotheses for its development.

Secondly, we look at how a heritage language develops in contact with a dominant one. Many children of immigrants in Australia grow up with language exposure that is 'situation-bound' (Vihman & McLaughlin 1982; Clyne 2003; Qi, Di Biase & Campbell 2006), where their two languages, a heritage language and English, develop separately (Meisel 1989; De Houwer 1990, 2005; Itani-Adams 2007). As far as Serbian is concerned, the Serbian input children receive is limited to home-based language use in communication with family members. Such a limitation in language input, coupled with the lack of opportunities for output (cf. Swain's 1995 output hypothesis) in a predominantly English-speaking environment, may affect the development and attainment (Doughty & Long 2003) of Serbian knowledge and skills. It is likely that the case-marking system will be affected because Serbian, which is a heavily morphologised case-marking language, is in contact with English, which is a highly configurational language with a much-reduced morphological expression of case. In testing our hypotheses on production data from three Serbian-Australian teenagers, all early bilinguals living in Sydney, we then expect to find

that their Serbian case system is not as fully developed as that of a comparable native speaker.

Thirdly, we look at how the case system can interface with pragmatic choices, in a similar way to Artoni & Magnani (cf. ch. 5, this volume). Because Serbian case markers on nominals identify functional roles (Hammond 2005: 105), they help give structural realisation to pragmatic factors such as topicality and prominence in the sentence (Browne 1993). Whilst the default Serbian word order is SVO, constituents are largely ordered by topic-comment structure. So, a limited availability of the case-marking system in their weaker language is important when discussing bilinguals' choices for structures at the interface between syntax and pragmatics. Thus we also use PT to assess how a more developed case-system can allow for a wider range of pragmatic-discourse driven syntactic choices.

2. The Serbian case system

A case system is a prominent characteristic of dependent marking languages (cf. (1), part II, this volume), and – as we have seen in chapter 5 about Russian – is traditionally defined, in a general way, as a system marking dependent nominals to the type of relation they bear to their heads in a phrase (Blake 1994). Ns can therefore depend on heads belonging to various lexical categories: verb, noun, adjective and preposition. Each of these lexical categories is associated with its typical cases for the Ns within their VP, NP, AP, and PP respectively. Among these lexical categories, V is crucial for the construction of the sentence because the cases it requires for its thematic roles identify their GFs. A corollary of this definition of case is that, when N is independent of any other words and fulfils the SUBJ function, it takes the basic 'unmarked' form of the nominative case.

There are seven cases in Serbian: nominative, accusative, locative, genitive, instrumental, dative and vocative. In this chapter, vocative will no longer be mentioned because it is independent of any other element in the sentence. Each lexical category governs its typical case, such as ACC for V, and GEN for N, as shown respectively in (1a-b).

- (1) a. mačka tera miša
 cat chase mouse-ACC
 [the cat is chasing a mouse]
- b. interpretacija romana je veoma interesantna
 interpretation novel-GEN is very interesting
 [interpretation of the novel is very interesting]

However, Serbian lexical categories may govern other cases, as shown in (2a-b), where the OBJ required by V is expressed by GEN and INST respectively rather than by ACC. Needless to say, these less systematic associations (or more marked cases) are harder to acquire.

- (2) a. baka je nakopala krompira
 grandma be dug up potatoes-GEN
 [grandma dug up the potatoes]
- b. kralj vlada zemljom
 king rule country-INST
 [the king rules the country]

Like verbs, also nouns, adjectives and prepositions can assign a variety of cases. For example, besides GEN, shown in (1b), N can assign DAT, as in (3).

- (3) spomenik Puškinu je jako visok
 monument Pushkin-DAT is very high
 [the monument to Pushkin is very high]

As case assigners, prepositions are especially taxing for the learner. Few of them, if any, select only one case; and few cases, if any, depend on only one preposition. So, on the one hand, displaying high functional overlap with their homonymy, polysemy, synonymy (Savić & Anđelković 2007), most prepositions select their case according to their different meanings. For example, the preposition *u* governs LOC when the meaning of PP is locational 'in', and ACC when it is directional 'to' as in (4a-b) respectively.

- (4) a. ja sam u sobi
 I am in room-LOC
 [I am in the room]
- b. idem u sobu
 am going to room-ACC
 [I am going into the room]

On the other hand, ACC for example, can be governed by *u* (directional 'to') as we have just seen in (4b), by *u* (temporal 'in') as in (5a), or *za* ('for') as in (5b).

- (5) a. u moje slobodno vreme pravim čestitke
 in my spare time-ACC draw cards
 [in my spare time I draw cards]

- b. prodajem kolače za svadbe i za krštenja
 sell cakes for weddings-ACC and for christenings-ACC
 [I sell cakes for weddings and for christenings]

Further complexities can arise from the semantics of certain NPs themselves, rather than from the element from which they depend. For example, Serbian NPs containing the numerals *two*, *both*, *three* or *four* obligatorily require an invariable paucal form (PAUC; Franks 1994: 606), and NPs containing a quantifier obligatorily require GEN. Thus, constituents functioning as SUBJ or OBJ may occur in this PAUC form as in (6a-b), or in GEN as in (6c-d). Likewise, NPs in PPs containing a quantifier occur in GEN regardless of the preposition, as in (7a-b).

- (6) a. troje dece čita u parku
 three children-PAUC.SUBJ read-3.SG in park
 [three children are reading in the park]
- b. imam troje dece
 have-1.SG three children-PAUC.OBJ
 [I have three children]
- c. mnogo pasa trči u parku
 many dogs-GEN.SUBJ run-3.SG in park
 [many dogs are running in the park]
- d. imam nekoliko hobija
 have-1.SG several hobbies-GEN.OBJ
 [I have several hobbies]
- (7) a. Sergej popravlja sto sa novim čekićem
 Sergej fix table with new hammer-INST
 [Sergej is fixing the table with a new hammer]
- b. Sergej radi sa nekoliko čekića
 Sergej work with several hammers-GEN
 [Sergej is working with several hammers]

Yet another set of difficulties for the learner is created by the fact that Serbian morphology, like Russian, is highly fusional. On N, for example, a single inflectional morpheme carries information about features such as gender (masculine, feminine or neuter), number (singular or plural), and case (nominative, accusative, genitive, dative, instrumental, or locative). Compounded with all this, there are three phonologically-based N classes: the first class, arising from Proto-Slavonic *o*-stems, includes most masculine and all neuter Ns ending in *-o*, *-a* or a consonant in NOM

SG and *-a* in GEN SG; the second class, continuing proto-Slavonic *a*-stems, includes most feminine Ns and some masculine Ns ending in *-a* in the NOM SG and *-e* in GEN SG; the third class, from Proto-Slavonic *i*-stems, includes all feminine Ns apart from *a*-stems, ending in a consonant in NOM SG and *-i* in GEN SG (Browne 1993: 319-322). No wonder, then, that the cognitive load required in sorting out form-to-function mapping (Pienemann 1998: 155) is heavy. Consider, for example, the case declensions shown in (11). With the *-o* stem N *grad* ('city'), NOM and ACC suffixes coincide in the singular, and so do DAT and LOC with *-u*. Also, both with masculine and feminine Ns, DAT, LOC and INSTR are practically identical in the plural: what distinguishes them in spoken production, however, is prosody (Browne 1993: 319).

(8) Declension by case of the masculine noun *grad* ('city') and the feminine noun *žena* ('woman, wife')

CASE	SINGULAR		PLURAL	
	MASC	FEM	MASC	FEM
NOM	grad	žena	gradovi	žene
ACC	grad	ženu	gradove	žene
GEN	grada	žene	gradova	žena
DAT	gradu	ženi	gradovima	ženama
LOC	gradu	ženi	gradovima	ženama
INST	gradom	ženom	gradovima	ženama

A final difficulty for the learner should be mentioned. In Serbian, within the NP nominal modifiers that precede the head N must agree with the head in gender, number and case (Hammond 2005). In other words, Serbian case must be computed formally not only on the head N or pronoun but also on adjectives and some quantifiers (Hammond 2005), as illustrated in (9). This requirement allows further flexibility in Serbian word order (not just the freeing up of SVO), in so far as elements within an NP can be discontinuous, as in (10) which may be used to express discourse-pragmatic requirements of the speaker.

(9) peglam ovu plavu haljinu
 iron-1.SG this-ACC blue-ACC dress-ACC
 [I iron this blue dress]

(10) ovu sam plavu haljinu peglala
 this-ACC be-1.SG.AUX blue-ACC dress-ACC ironed
 [I have ironed this blue dress]

In our analysis here, we will deal mainly with case construction by V. Because of the pivotal role V plays in building up the sentence (cf. § 2.1, ch. 1, this volume), we are crucially interested in how case can mark the GFs of their arguments (i.e., SUBJ, OBJ, and OBL_θ) and less so in the internal structure of NPs, which would require a larger study and more space than is available here. Among GFs, however, we ignore COMP because, although argumental, it is unmarked by case as such (although it may of course contain case in its internal structure). On the other hand, because OBL_θ are most often expressed by PPs, we take into consideration also prepositions as case assigners. In other words, here we focus on how learners first learn to associate a set of lemma specifications with a set of GFs and build up their f-structure by means of case markers, and then exploit their functionally case-marked sentence constituents for discourse and pragmatic reasons. Let us then look at cases as markers of GFs.

According to Nordlinger's (1998) quadripartite typological scheme reported in (1) in part II, this volume, Serbian is an example of a nonconfigurational dependent-marking language. This means that, on the one hand, when linking the lexicon to c-structure, along the configurational continuum, Serbian f-structure information is expressed by morphology (rather than by position, as in English). On the other hand, along the dependency continuum, Serbian GFs are marked inflectionally on the depending element (rather than on the head, as in Italian, cf. ch. 3, this volume). Marking a dependent element inflectionally means using case. As a matter of fact, as remarked in chapter 5 about Russian, case is one of three devices by which, typologically, languages can identify GFs, especially the core ones SUBJ and OBJ – the other two devices being word order as in English, and agreement as in Italian (Kroeger 2005: 102-ff). In brief, Serbian relies on obligatory N morphology for identifying GFs, regardless of word order.

Case as function assigner (together with SUBJ-V agreement, which in Serbian is a further device identifying core grammatical relations) allows for great flexibility in the word order of the Serbian sentence. So, besides canonical SVO order, all the other five permutations of the three core elements in a sentence are grammatically acceptable in Serbian: SOV, VSO, VOS, OSV, OVS. These six orders do not exhaust all possible sequences because all of them are possible with null SUBJ and null OBJ. And speakers exploit them all to organise sentences according to the pragmatic requirements of the topic-focus structure. Once constituents are marked functionally by case, they can be positioned varyingly in the sentence. For example, the same propositional content, expressed with SVO in a pragmatically unmarked way in (11a), can be expressed with different word order if the speaker wishes to topicalise OBJ in (11b).

- (11) a. mačka tera miša
 cat-NOM.SUBJ chase-3.SG mouse-ACC.OBJ
 [the cat is chasing a mouse]
- b. miša tera mačka
 mouse-ACC.OBJ chase-3.SG cat-NOM.SUBJ
 [the cat is chasing a mouse / it is a mouse that the cat is chasing]

In sum, for learners of Serbian, case is a complex feature to acquire for a variety of reasons: morphologically, there are several cases, fusionally enmeshed with other nominal features such as number, gender and class. Morphosyntactically, case must be computed on most nominal elements within the NP. Syntactically, cases identify GFs in the sentence. Most of the time, relations between case and function are default and predictable. However, less predictably, the same case can construct alternative GFs, and the same GF can be constructed by different cases, although lexically with different predicates. On the basis of these complexities, we now suggest some hypotheses for the development of the case system in learners of Serbian as a heritage language.

3. The developmental hypotheses

Morphologically, we hypothesize that, among the Serbian cases, the first to be used will be NOM for three main reasons: NOM is the citation form; it often coincides with the ACC form in the singular; and it is the case learners pervasively find in the prominent first position in the sentence. Once they notice variation in form, early contrasts may be set up as NOM-nonNOM with the nonNOM form(s) overextended. Then the sequencing in the spread of case forms from the emergence of a first contrast to the whole paradigm is an empirical matter, which may be subject to variation, both contextual and individual. And it goes without saying that we are talking here about the emergence of case-markers, without any consideration for their formal accuracy in terms of gender, number, or class.

Morphosyntactically, we predict that case will be marked according to PT's well-tested hierarchy based on the activation of the processing procedures which allow for the exchange of information about case first within the phrase, and then within the sentence. As already mentioned, the focus in our analysis here is at sentence level, i.e., on cases marking GFs, so we will not take into account case agreement phenomena. On the other hand, the case marker of the N in a PP bearing the OBL function will be considered, even if this is specified in the lexical entry of the preposition, and hence phrasal in terms of morphological development and independent of functional assignment in terms of syntactic development.

Syntactically, the sequence in (12) shows our developmental hypothesis for Serbian declaratives based on PT's Prominence Hypothesis (cf. § 4.2.1, ch. 1, this volume, for the universal sequence, and chh. 2, 3 and 4 for English, Italian, and Japanese respectively). At the canonical word order stage, the first match will occur between form and position; thus in an SVO language like Serbian, NOM is associated with preverbal position, and ACC with postverbal position. As a matter of fact, to our English Serbian bilinguals, case marking will actually seem redundant (as it may have happened historically for English). Only later, when functional assignment is in place, will learners be able to match NOM with SUBJ, ACC with OBJ, and DAT with OBL independently from the positionally determined SVO blocked sequence. So, at the noncanonical word order stage, they will be able to depart from the rigidity of the fixed SVO sequence in order to express their own discourse and pragmatic choices.

(12) *Developmental stages hypothesised for Serbian L2 syntax based on the Prominence Hypothesis: declaratives*

STAGE	STRUCTURE	EXAMPLE
NONCANONICAL WORD ORDER	OV(S) (topicalisation of OBJ)	<i>supu voli devojčica</i> [soup-ACC.OBJ likes girl-NOM.SUBJ] <i>supu devojčica voli</i> [soup-ACC.OBJ girl-NOM.SUBJ likes]
	VS(O) (focalisation of SUBJ)	<i>voli devojčica supu</i> [likes girl-NOM.SUBJ soup-ACC.OBJ]
XP CANONICAL WORD ORDER	TOP SVO (topicalisation of ADJ)	<i>svakog dana đaci imaju ispit</i> [every day-GEN.ADJ students-NOM.SUBJ have exam-ACC.OBJ]
CANONICAL WORD ORDER	SVO	<i>jedem krofnu</i> [(I) eat doughnut-ACC.OBJ] <i>devojčica voli supu</i> [the girl-NOM.SUBJ likes soup-ACC.OBJ]
LEMMA ACCESS	single words formulas	<i>ja ser</i> [I (like) cheese] <i>zovem se Mary</i> [my name is Mary]

The reader should bear in mind that, like the previous study on Russian, this is an exploratory study of case – not only for Serbian but also within PT's framework. Thus not all hypotheses are testable on our current cross-sectional data. Assuming that teenage Serbian-English bilinguals in Australia have positive attitudes towards preserving their mother tongue and use it in the home environment to communi-

cate with family and friends, it is expected that they would use native adult-like skills in sentence processing, should they possess them. Using as a benchmark the presumably full Serbian case system of our adult bilingual informant with schooling in Serbia, we therefore focus on the range of structures and case markers displayed by our three bilingual teenagers, and the way they are deployed to allow for speaker perspective and expressiveness beyond canonical word order rigidity.

4. Method

As already mentioned in § 1, the informants in this study are three teenage Serbian-Australian bilinguals. All three acquired both languages as first languages, with Serbian as the home and community language, and English as the dominant one. Their families originate from Vojvodina, a northern region of Serbia, so all of them speak the same Ekavic dialect. A Serbian native speaker who migrated to Australia when she was 20, and thus learned English as an adult, also participated in this study to ensure that the range of targeted structures is actually produced by native speakers in similar situations of elicitation. She too originates from the same Serbian region. Further details about the four informants are shown in (13).

(13) *The informants*

CODE-NAME	AGE	COUNTRY OF BIRTH	AGE AT ARRIVAL IN AUSTRALIA	OTHER INFORMATION
TRISH	13	Australia		no instruction in Serbian
NED	15	Australia		no instruction in Serbian
DON	17	Serbia	1;5	2 years of Serbian classes
NICOLE	35	Serbia	20	in Australia for 15 years

Data elicitation took place in dyadic conversations between the informant and the Serbian researcher. Informants were asked to perform three communicative tasks. During the short conversation, they responded to questions about their personal experience for the purpose of gathering information on language maintenance strategies. During the spot-the-difference task, they could report on at least seven differences between the two pictures shown. In the story-telling task, they narrated the children's story of 'Goldilocks and the three bears' following a sequence of pictures.

Among the whole corpus, the data set analysed for this study consists of 231 Serbian sentences produced by our four bilingual speakers. In accordance with the current state of PT in general, and its Prominence Hypothesis in particular, these are all main declarative sentences with finite lexical V. This means that we do not

consider constructions involving subordination, as well as reflexives, passives, and causatives. Also copulas and presentatives are excluded because they are ‘non verbal predicates’ (Kroeger 2005: ch. 10). Of these 231 sentences, 40 are irrelevant to our aims here because they are unmarked for case, that is, they bear only GFs which are COMP, OBLLOC or ADJ expressed with adverbs as in (14), or have all participants express by a codeswitch from English as in (15).

- (14) a. Ned sad je htela da spava
 now-ADJ be-3.SG.AUX want-3.SG to sleep-COMP
 [now she wanted to sleep]
- b. Ned ušli su unutra
 entered be-3.PL.AUX inside-OBLLOC
 [they went inside]
- (15) Tri Charmed ima one witchy sort of stuff
 Charmed-Ø.SUBJ have-3.SG that witchy sort of stuff-Ø.OBJ
 [Charmed (television show) has that witchy sort of stuff]

Thus our main analysis in § 5 considers 191 sentences, all of which mark at least one GF by case.

5. The analysis

In (16) we show the distribution of syntactic structures among the four informants. In order to avoid cluttering up a single table, we consider first only argument functions, and then deal separately with the various positions of the nonargument function ADJ. This explains why the XP canonical word order stage with its typical ADJ SVO structure is missing in (16).

As expected, by far the most frequent way of organising syntax is by canonical word order with all the informants, although it is worth noticing that all three early bilinguals dominant in English tend to avoid the more frequent null-SUBJ sentences typical of Serbian monolingual speakers, and use overt pronominal or referential SUBJ more often than the late bilingual Nicole does – especially Trish, who uses this structure more than twice as much. Pragmatically marked structures are much rarer, both those that place GFs other than SUBJ in preverbal position and those that place SUBJ in postverbal position for lexical requirements.

Regarding noncanonical word orders, we observe several notable differences between our informants. Foremost is the fact that Nicole’s range is much wider than that of the other three bilingual speakers. For example, the sentences in (17)

(16) *Cross-sectional study: distribution of argument functions among structures and informants*

STAGE	STRUCTURE	TRI	NED	DON	NIC	TOTAL
NONCANONICAL WORD ORDER	OBJ V SUBJ				1	1
	V SUBJ OBL			4		4
	V SUBJ				4	4
	OBJ SUBJ V				1	1
	OBJ V		1	2	1	4
	OBL SUBJ V				2	2
	OBL V				1	1
	SUBJ OBJ V	1				1
	SUBJ V OBL _{CASE}				1	1
	SUBJ V OBL _{PP}	4	2	2	4	12
CANONICAL WORD ORDER	SUBJ V OBJ	15	6	9	8	38
	SUBJ V	13	10	9	5	37
	V OBL	8	3	7	13	31
	V OBJ	14	15	10	15	54
	TOTAL	55	37	43	56	191

illustrate clearly how Nicole is the only informant producing postverbal SUBJ in the context ('the bear said/replied/asked') of the Goldilocks story retelling task. In this context even Don avoids postverbal SUBJ, although he is clearly able to use it with unaccusative Vs, as in (18).

- (17) a. Nic odgovorio je mali medved
 replied be-3.SG.AUX small bear-NOM.SUBJ
 [the small bear has replied]
- b. Don mali medved kaže
 small bear-NOM.SUBJ say-3.SG
 [the small bear says]
- c. Ned najmanji medved je rekao
 smallest bear-NOM.SUBJ be-3.SG.AUX said
 [the smallest bear has said]
- d. Tri mali medveć¹ je pito
 little bear-NOM.SUBJ be-3.SG.AUX asked
 [the little bear has asked]

1 *Medveć* is the informant's version of the diminutive *medvedić*. Case, however, is used appropriately.

- (18) Don došo muški medved do svoje sup
 come-3.SG male bear-NOM.SUBJ next to his soup-OBL
 [the male bear came towards his soup]

Furthermore, Nicole is the only informant who topicalises OBL, as shown in (19).

- (19) Nic na mojoj stolici je neko sedeo
 on my chair-LOC.OBL be-3.SG.AUX someone-NOM.SUBJ sat
 [someone has sat on my chair]

Finally, Ned and Trish use only one noncanonical structure each, neither of which involves postverbal SUBJ, as shown in (20). On the other hand, Nicole and Don use four and five postverbal SUBJ structures respectively, as already exemplified in (17a) and (18).

- (20) a. Ned istoriju znaš volim
 history-ACC.OBJ (you) know like-1.SG
 [history, you know, I like]
- b. Tri ona je to sve pojela
 she-NOM.SUBJ be-3.SG.AUX that all-ACC ate
 [she ate all that]

Let us now look at the position of ADJ in the production of our informants. In Serbian there are no constraints on where it can be placed, and speakers are free to place it according to their discourse or pragmatic need. However, the three positions ADJ can occupy with reference to canonical word order (after, within, and before) gain different significance in terms of developmental syntax. If ADJ follows canonical word order, it belongs to the early syntactic stage, which clearly also Trish has safely reached. If ADJ occurs between the core GFs, namely between either SUBJ and V, or V and OBJ, it is a sign that learners are starting to free up the SVO block, and GFs are somehow no longer assigned by position alone. However, in this case ADJ does not compete with SUBJ for the association with the initial DF TOP. Also when ADJ is before V in a sentence with null SUBJ it is not, by itself, sufficient evidence that a learner has reached the XP canonical word order stage. Since this latter stage involves the separation of SUBJ and TOP, we assume that the passage of stage is clear only when SUBJ is present and competes for prominence with ADJ as TOP in sentence-initial position.² So, if ADJ precedes SUBJ, there is a more sub-

² In § 3.1, ch. 3, this volume, Di Biase & Bettoni place the TOPADJ VO structure at the higher XP canonical word order stage. Be as it may, unequivocal evidence for the passage of stage is provided when SUBJ is present and competes with ADJ as TOP.

stantial sign of having reached the XP canonical word order stage. Finally, any other position of ADJ occurring with marked orders is no longer developmentally significant, because by then learners have freed up fully the more important core GFs. In (21) we show where our informants stand with regard to the placement of ADJ in the 64 sentences containing this GF among the total 191 ones.

(21) *Cross-sectional study: distribution of ADJ among structures and informants*

STAGE	STRUCTURE	TRI	NED	DON	NIC	TOTAL
NONCANONICAL WORD ORDER	OBL ADJ V				1	1
	OBJ ADJ V			1		1
XP CANONICAL WORD ORDER	ADJ SUBJ V X		2	2	1	5
CANONICAL WORD ORDER	ADJ V X	2	3	2	6	13
	SUBJ ADJ V X	3	1	3	3	10
	(SUBJ) V ADJ X	2	1	3	2	8
	(SUBJ) V X ADJ	4	5	9	8	26
TOTAL		11	12	20	21	64

In the table we can see again how, compared to Trish and Ned, Nicole and Don use a wider range of positions, including at least one example each of topicalised ADJ with referential or pronominal SUBJ. Ned too has two examples of this structure, as in (22), but Trish has none, although she can vary the ADJ position within canonical word order as in (23a-b), and can topicalise ADJ in a null-SUBJ sentence as in (23c). We explain Trish's and Ned's smaller range of structures by the fact that their functional assignment relies more on position than on case.

- (22) Ned prvo je sela na najveću stolicu
 first be-3.SG.AUX sat on biggest chair
 [first (she) sat on the biggest chair]

- (23) Tri a. moja samo ima yellow suknju
 my-NOM only-ADJ have-3.SG yellow skirt-ACC
 [my (girl) only has a yellow skirt]

- b. ja volim malo one lolies
 I-NOM like-1.SG a little bit-ADJ those lollies-Ø
 [I like those lollies a little bit]

- c. samo vidiš jednu glavu
 only-ADJ see-2.SG one head-ACC
 [you only see one head]

In (24) we show how our four informants can handle case marking of direct (SUBJ and OBJ) and indirect (OBL_{CASE} and OBL_{PP}) argument functions. Not included in the analysis is case marking in ADJ, even when these are expressed by PPs rather than adverbs because it would not provide more information about the use of case than OBL already does when it too is expressed by a PP. Furthermore, cases are counted with a + sign when appropriately used in a given context, whereas the brackets around a case after the figure preceded by the - sign indicate which case is used inappropriately. Appropriateness, however, should not be confused here with accuracy (Pienemann 1998). That is, keeping in mind that, as we have already mentioned in § 2, a single inflectional morpheme in Serbian may be used for several morphological contrasts (e.g., number and gender, as well as case), we accept as appropriate any target case ending regardless of gender, number etc. When these are hard to factor out, we give the speaker the benefit of the doubt.

(24) *Cross-sectional study: distribution of case markers in relation to argument functions among informants*

FUNCTION	TARGETED CASE	TRI	NED	DON	NIC
OBL_{PP}	prep INST				+1
	prep LOC	-3(ACC)	-2(ACC)	+1	+6
	prep GEN			+4	+3
	prep ACC	+5 -1(DAT)	+3	+4 -3(LOC)	+9
	total OBL_{PP}	+5 -4	+3 -2	+9 -3	+19
OBL_{CASE}	DAT				+1
	total OBL_{CASE}				+1
OBJ	Ø (codeswitch)	+2	+1		
	PAUC (numeral)	+4	+7	+4	+5
	GEN (quantifier)				+2
	ACC	+22 -1(NOM) -1(GEN)	+14	+17	+19
	total OBJ	+28 -2	+22	+21	+26
SUBJ	PAUC (numeral)		+2	+2	
	GEN (quantifier)			+1	
	NOM	+33	+16	+22	+26
	total SUBJ	+33	+18	+25	+26

Predictably, Nicole's case marking system is fully accurate. Furthermore, her range of case use is wider than that of the other informants, including also INST and DAT. Whereas INST occurs within a PP, she is the only informant who uses DAT in order to mark an oblique argumental function (OBL_{CASE}), as shown in

(25), rather than marking it by a OBL_{pp} . As a matter of fact, also other informants mark OBL_{LOC} by means of DAT more than once, as in *tata osto kući* ('father stayed home'), but we consider this frequent use of *kući* formulaic, and ignore it in our analysis.

- (25) Nic mama medvedica je rekla medvedu
 mother bear-NOM.SUBJ be-3.SG.AUX said bear-DAT.OBL
 [mother bear has told the bear]

All other informants show some uncertainty in the use of case markers. With Don and Ned, errors in case marking are limited to PPs, and thus attributed to the lexical entry of individual prepositions, with little relevance for functional assignment, and consequently minimal risk for misunderstanding, in so far as the presence of the preposition facilitates comprehension of the intended meaning (Anđelković 2000); an example is shown in (26), where in OBL_{LOC} Ned uses ACC instead of LOC .

- (26) Ned neko je sedeo na njegovu *stolicu
 someone-NOM.SUBJ be-3.SG.AUX sat on his chair-ACC.OBL
 [someone has sat on his chair]

With Trish, however, inaccurate case marking affects also core arguments. Although all her $SUBJ$ s are marked as NOM , as (24) shows, and her range of $NOM-ACC$ distinctions is quite large, including both nominal and pronominal elements, as shown in (27), she marks OBJ twice with a wrong case: once with NOM and the other with GEN , as shown in (28a-b respectively).

- (27) Tri a. veći medveć spasio je onu malu
 bigger bear-NOM.SUBJ saved be-3.SG.AUX that little-ACC.OBJ
 [the bigger bear has saved that little one (girl)]
- b. ona je probala onu treću
 she-NOM.SUBJ be-3.SG.AUX tried that third-ACC.OBJ
 [she has tried that third (soup)]
- (28) Tri a. ona samo drži mali *ker
 she-NOM.SUBJ only hold-3.SG.AUX little dog-NOM.OBJ
 [she is only holding a little dog]
- b. onak su videli *kreveta
 then be-3.SG.AUX see-3.SG.AUX bed-GEN.OBJ
 [then they saw beds]

This uncertainty in case marking displayed by Trish is also evident in several PPs, where she tends to oversupply ACC. As a matter of fact, it would seem that Trish is still at the stage of NOM vs nonNOM variation along the morphological schedule, and that her nonNOM forms only happen to coincide with ACC, DAT and GEN forms in the native language – as suggested by one of the anonymous referees of this chapter. Notice also that Trish is the informant who most often leaves constituents unmarked by case by using English, as shown in (29). The use of this type of codeswitching from the dominant language into a heritage language is a well attested characteristic of first and second generation migrants' speech in general, in Australia (Bettoni 1991; Clyne 2003) as elsewhere (Auer 1998; Li Wei 2006). But there can be no doubt that the more frequent the codeswitches, the weaker the case system (cf. Schachter's (1974) avoidance strategy).

- (29) Tri ja volim visual arts
 I-NOM.SUBJ like-1.SG visual arts- Ø.OBJ
 [I like visual arts]

Even if not altogether apparent in the tables and examples presented so far, the 191 sentences analysed for this study provide us with further evidence that the Australian young bilinguals' case marking system is less reliable than Nicole's, who is a late bilingual with Serbian schooling. For example, compare Nicole's discontinuous OBJ constituent in her OSV sentence in (30a) with Don's SVO canonical sequence in uttering the very same referential content in (30b). Even though Don's sentence is grammatically correct, there can be no doubt that it lacks expressiveness compared to Nicole's.

- (30) a. Nic moju je supu neko pojeo celu
 my-ACC be-3.SG.AUX soup-ACC.OBJ someone-NOM.SUBJ eaten entire-ACC
 [someone has eaten my entire soup]
- b. Don neko je pojeo celu moju supu
 someone-NOM.SUBJ be-3.SG.AUX eaten whole-ACC my-ACC soup-ACC.OBJ
 [someone has eaten my whole soup]

One could argue that English interference can explain both reliance on SVO and caseless transfers. Yet, in our data there are other structures where all three Australian bilinguals seem to have problems despite a similarity between their two languages. A case in point is possessive GEN as a marker of dependency from N. Don produces three sentences in which GEN correctly marks the possessor (as the English morpheme *-s* does), but sometimes insecurity with the case lets him add the preposition the preposition *od*, as in (31a), which turns out to be ungrammat-

ical because an N+N dependency would be expected by native speakers in this context. Ned seems to ignore the issue by marking all elements in the OBJ constituent with ACC in (31b), thus missing the N+N dependency. And Trish seems to lose track of her sentence by using a preposition and then marking *little (girl)* with ACC instead of GEN in (31c).

- (31) a. Don na kraju je išla od malog deteta na *stolici
 in end-LOC.ADJ be-3.SG.AUX gone of little child-GEN on chair-LOC.OBL
 target: na kraju je išla na stolicu malog deteta
 in end-LOC.ADJ be-3.SG.AUX gone on chair-ACC.OBL little-GEN child-GEN
 [in the end she went to the little child's chair]
- b. Ned jedna ima *ljubičastu *boju odelo
 one-NOM.SUBJ have-3.SG.AUX purple-ACC colour-ACC outfit-ACC.OBJ
 target: jedna ima odelo ljubičaste boje
 one-NOM.SUBJ have-3.SG.AUX outfit-ACC.OBJ purple-GEN colour-GEN
 [one (woman) has a purple coloured outfit]
- c. Tri njena majka, od te Goldilocks *malu, je došla
 her mother-NOM.SUBJ of that Goldilocks little-ACC (girl) be-3.SG.AUX came
 target: njena majka, majka male Goldilocks, je došla
 her mother-NOM.SUBJ, mother-NOM little-GEN Goldilocks, be-3.SG.AUX came
 [her mother, little Goldilocks' mother, came]

6. Conclusion

In this chapter we have dealt with a case marking system within the PT framework, and considered both morphological and syntactic aspects of its development. Our example of the Serbian language is particularly challenging for the learner (and the researcher), not only because, morphologically, there are several cases which mark N dependency on several lexical categories and whose formal features are fusionally 'mixed' with those of gender, number and class features, but also because, besides a more systematic, productive use of default case assigned syntactically, there are numerous instances of lexically triggered nondefault case which override syntactic assignment and must be learned individually, or at best bundled up in small sets. We have then offered some hypotheses for the development of the Serbian case marking system, both morphological and syntactic, and tested them on cross-sectional data pertaining to four bilingual informants whose knowledge of Serbian is already fairly advanced.

Our analysis of the data available cannot provide evidence for the full developmental path that lies ahead for learners acquiring Serbian as a heritage language.

Nevertheless, they confirm several of our PT-based hypotheses. First, syntactically, no informant producing structures at the noncanonical word order stage has not already safely in place those at the canonical word order stage, sequencing verbal arguments in canonical order. Secondly, the position of the nonargument function ADJ, freely placed according to discourse or pragmatic reasons, is the playground for allowing learners to free up the default canonical word order. Thirdly, no informant who displays a full range and accurate use of case markers does not exploit the possibility of deploying them in order to produce noncanonical word orders. On the other hand, the opposite seems to hold true. Our informants whose case system still shows some gaps or formal inaccuracies rely on the rigidity of position within canonical word order to identify GFs, or else appeal to the semantically more transparent PPs.

In sum, focusing on how V constructs the relationship with its arguments by means of case, PT has allowed us to show how a wider range and a more reliable deployment of formal cases correlates with a stronger possibility of exploiting them to enhance discourse and pragmatic choices beyond canonical word order, thus allowing expressiveness for the speaker without compromising comprehension by the listener. This in turn has allowed us to discriminate between advanced and less advanced heritage speakers of Serbian, and propose a new approach to the analysis of a minority language. In our approach, contrary to Dimitrijević's (2004a, 2004b; Dimitrijević-Savić 2008), the contact between the two typologically different languages (i.e., nonconfigurational Serbian and configurational English) plays a lesser role than immature development, which is constrained by a migration environment offering 'situation-bound' language exposure and arguably a much-reduced input in the heritage language, both in quantity and quality. Further research focusing, for instance, on the comparative development of both the minority language and the dominant language once the child begins his/her school experience may clarify the role of the specific sort of input afforded by instruction (but cf. Medojević 2014). We leave it to more robust data – gathered longitudinally or cross-sectionally among more numerous learners with more varied competence levels – to test the full range of developmental hypotheses concerning the acquisition of the case marking system of Serbian (and other languages), and thus confirm, or falsify, our claims.

The authors wish to thank Rosanna Benacchio (University of Padua, Italy) and Lidija Krebs-Lazendić (University of New South Wales, Australia) for their helpful suggestions and proofreading of the Serbian language.

Exploring the acquisition of differential object marking (DOM) in Spanish as a second language

Bruno Di Biase* and Barbara Hinger**

*University of Western Sydney, **University of Innsbruck

1. Introduction

The aim of this chapter is to explore the acquisition of differential object marking (DOM) in Spanish L2 and thus probe the higher level boundaries of the PT framework. In coining the term DOM, Bossong (1983-84, 1991) presented cross-linguistic data on more than 300 languages presenting this grammatical characteristic, whereby direct objects (OBJs) of transitive Vs either remain unmarked or are overtly marked by case or agreement on the basis of some semantic or pragmatic feature. This marking has since attracted considerable attention in linguistic theory (e.g., Aissen 2003; Dalrymple & Nikolaeva 2011; Leonetti 2004; Torrego 1998, 1999, among many others). Unlike the many purely structural approaches, Dalrymple & Nikolaeva's (2011: 1-2) point out that DOM, in the many languages where it manifests itself, "encompasses syntactic, semantic and informational-structural differences between marked and unmarked objects". So they propose that marked OBJs are associated, synchronically or historically, with the information-structure role of *topic*. Where the connection between marked OBJs and topicality has been lost through grammaticalisation, marked OBJs become associated with *semantic features* typical of topics, such as animacy, definiteness and specificity (Dalrymple & Nikolaeva 2011:1-2). Spanish is one of the languages exhibiting DOM, whereby some OBJs are marked with the preposition *a*, also known as marked accusative (Torrego 1998), prepositional accusative, personal *a*, or accusative *a* (Montrul & Bowles 2008; Tippets 2011).

Given our interest in acquisition, we note that Spanish DOM is highlighted in the literature as difficult to acquire, not only for English L1 learners of Spanish L2 (Bowles & Montrul 2009; Guijarro-Fuentes 2011, 2012; VanPatten & Cadierno 1993), but also, perhaps more surprisingly, for early bilingual 'heritage' speakers of Spanish in the USA (Montrul 2008; Montrul & Bowles 2009; Montrul & Sanchez-Walker 2013; Silva-Corvalan 1994). For L2 Spanish, Farley & McCollam

(2004) confirm this difficulty for *a*-marked OBJs. Their findings however show no support for the PT-based schedule they derive from Johnston (1995) and Pienemann (1998), who place *a*-marked OBJs well before subjunctives, which are at the top stage in PT. To our mind this is not surprising because the earlier version of PT locates structures on a single developmental axis, whether they are obligatory or optional, whether declaratives, negatives, interrogatives, or other pragmatically motivated constructions. Thus earlier PT is not equipped to deal with optional and interface phenomena, including Spanish DOM. On the other hand, the current proposal for PT by Bettoni & Di Biase in chapter 1 of this volume offers principled explanations and more attuned predictions in the area of syntax-semantic and syntax-discourse interfaces.

This chapter then sallies into an exploration of the rather controversial area of case marking, itself a relatively new area in PT, but limited to DOM. More specifically it will attempt to show that the current version of PT is better suited to account for the difficulties in learning differential case-marking, which we see as a structure located at the interface of syntax-semantics and syntax-discourse. Similar observations about the difficulties advanced adult L2 learners face when acquiring the interface of syntax with other cognitive domains led Sorace & Filiaci (2006) to propose the Interface Hypothesis, attributing the difficulty, possibly, to computational limitations in integrating multiple sources of information. These difficulties have been variously interpreted also in terms of 'incomplete acquisition' (e.g., Montrul 2008), from lack of access or partial access to universal grammar, or some kind of 'representational deficit' (Clahsen & Felser 2006), and in terms of the Feature (In)accessibility Hypothesis (Guijarro-Fuentes 2011, 2012).

We interpret these difficulties in processing terms, in the sense that the computational complexity created by the requirement to integrate discourse information with syntactic information makes processing in real time harder for the learner (Hopp 2007), as opposed to the native speaker, who has already automatized the necessary underlying processes. It is also plausible, as Wilson (2009) claims, that the additional computational complexity created by the attempt to integrate different layers of information challenges the ability of the learner to allocate cognitive resources appropriately. For example, competing constraints from the L1 may interfere on how to interpret the semantic or discourse requirements. Indeed both resource limitation and resource allocation may contribute simultaneously to the learner's difficulties. We will not delve any deeper into our processing option versus representational deficits as an explanation of the learners' difficulty with interface structures. We will instead suggest to place Spanish DOM, a structure sitting at the interface between syntax and semantics/discourse, within a current PT-based schedule for L2 Spanish, and propose some initial empirical tests for our position. In the remainder of this chapter we will first offer a quick sketch of Spanish and

the intriguing nature of its DOM, followed by a brief review of the acquisitional literature (§ 2). Next, we will present our developmental schedules for L2 Spanish over which DOM is distributed (§ 3), and empirically test them in a small-scale cross-sectional study of oral production of six Austrian students of Spanish as a foreign language (§ 4).

2. Spanish and its Differential Object Marking

According to Tippets (2011: 107), Spanish uses the preposition *a* to mark “human accusative (direct) objects. This transparent example of DOM relates to the animacy status of the accusative object”. The facts of Spanish, however, also support the proposition that neither all animate OBJs are *a*-marked nor that all inanimate OBJs are not *a*-marked. But before we zoom into this intriguing grammatical area let us zoom out to a brief overview of the language.

A pluricentric language, Spanish is by far the most widely spoken of the Romance languages, and the national language of 18 countries as well as Spain, with large Spanish-speaking minorities in the USA, and significant minorities in a number of other countries in the world (cf. Green 2011 for an accessible overview of this language). Spanish is also, naturally, subject to regional and sociolinguistic variation. It shares many characteristics with other Romance languages (cf. Italian, ch. 3, this volume) including nonconfigurationality, null SUBJ phenomena, a rich agreement morphology and pronominal clitics. Following Green’s (2011) description and modelling on his examples, Spanish word order is not fixed by grammatical requirements at a particular point in the sentence, which differentiates it from French and even more from the Germanic languages. But it has strong constraints *within* the main syntactic constituents, and its theoretical word order freedom is subject to pragmatic conventions: themes precede rhemes and new information is placed towards the end of the utterance.

In canonical word order, OBJs and complements follow V, as in (1a-b). Given the tendency for SUBJ and TOP to coincide in spoken language, the SVO order in (1a) is very frequent, whereas SOV in (1b) is register-dependent (e.g., in poetry for rhyming reasons); VOS in (1c) would sound very odd, and VSO (1d), although common in more formal registers, would signal contradiction or contrast in everyday language (‘it was Juana, not Carmen, who painted a car’).

- (1) a. Juana pintó un coche
 Juana-SUBJ painted-3.SG a car-OBJ
- b. Juana un coche pintó
 Juana-SUBJ a car-OBJ painted-3.SG

- c. pintó un coche Juana
 painted-3.SG a car-OBJ Juana-SUBJ
- d. pintó Juana un coche
 painted-3.SG Juana-SUBJ a car-OBJ

Similarly to Italian (cf. § 3.1, ch. 3), OBJ (when definite NP or proper N) can be topicalised by placing it at the beginning of the sentence, as in (2a), with an intonation break after TOP and an ACC clitic obligatorily marked on V.¹ As Green (2011) notes, the result of this ‘clitic copying’ is no longer a simple sentence but a complete and perfectly grammatical structure in its own right. In other words, an XP (*los coches/the cars*) is added to a complete S, which here comprises a transitive V with its OBJ clitic marker coreferential with the external XP, followed by SUBJ. Another important point is that, when OBJ is topicalised, SUBJ is postverbal as in (2a), assuming a secondary² TOP role. In any case the sentence is complete even when the overt TOP is dropped (2b) because the obligatory OBJ is supplied by the clitic (*los*) displaying anaphoric agreement with TOP.

- (2) a. los coches, los pintó Juana
 the cars-PL.MASC they-ACC.PL.MASC painted-3.SG Joan
- b. los pintó Juana
 they-ACC.PL.MASC painted-3.SG Joan

SUBJ is obligatorily postverbal also in presentationals (3a) and in content questions (3b). However, again following Green (2011), interrogatives should not be assumed to entail SV inversion because postverbal SUBJ frequently occurs in statements,³ and polar questions may show VS or SV order and rely on intonation to differentiate from statements.

- (3) a. hay muchos puentes en Sydney
 there are many-PL.MASC bridges-PL.MASC in Sydney

1 Like Italian, Spanish is a head-marking language, which therefore can mark OBJ morphologically on V (i.e., the head).

2 As Dalrymple & Nikolaeva (2011: 53–4) remind us, “(t)he topic role is not necessarily unique”. Along with other scholars (e.g., Givón 1983; Lambrecht 1994; Polinsky 1995), they distinguish at least primary topic and secondary topic.

3 A class of unaccusatives Vs with SUBJ=FOC normally exhibit a VS order, similarly to Italian (cf. § 3.2, ch. 1, this book): in the sentence *llegó el jefe* (‘arrived the boss’), postverbal SUBJ is the unmarked position.

- b. qué pintó Juana ?
 what painted-3.SG Joan ?

As for morphology, Spanish explicitly and consistently marks number and gender on all modifiers within NP, as well as number and person, and occasionally also gender, between SUBJ and V, as (4a) shows. This provides optimal grounds for testing classic PT. Interestingly for our concern in this chapter, Spanish – unlike Italian (cf. § 2.1, ch. 3, this volume) – has no anaphoric agreement between lexical V and OBJ, not even in constructions topicalising OBJ, as is confirmed by the grammaticality of (4b), where *comido* (eaten) does not carry the plural and feminine features of *las manzanas* (the apples).

- (4) a. las manzanas están maduras
 the apples-PL.FEM are 3.PL ripe-PL.FEM
- b. las manzanas las han comido los monos
 the apples-PL.FEM they-ACC.PL.FEM have eaten the monkeys-PL.MASC
 [the apples have been eaten by the monkeys]

Possible ambiguities in distinguishing between SUBJ and OBJ are resolved by syntactic differences in two important ways, both connected to specificity (Green 2011). The first syntactic difference is that the Spanish SUBJ NP – whether definite, indefinite or generic – requires a determiner, whereas the OBJ does not. This is shown by the grammaticality of (5a), where the SUBJ *el hombre* ('the man') appears with the determiner, and the OBJ *manzanas* ('apples') without it. The OBJ without the article (*manzanas*, 'apples') responds to the question 'what did the men buy?', and is hence a characteristic nonTOP, focused OBJ. On the other hand, (5b) responds to the question 'what happened?'. Hence the OBJ *las manzanas* ('the apples') is again a nonTOP, but neither is it FOC because, as 'event reporting' (Lambrecht 1994), the whole sentence is in focus (new information). With regard to SUBJ, (5c) is ungrammatical because the SUBJ *hombres* ('men') is without a determiner, even though it is generic.

- (5) a. el hombre compró manzanas
 the man bought-3.SG apples
- b. el hombre compró las manzanas
 the man bought-3.SG the apples
- c. *hombres compraron manzanas
 men bought-3.PL apples

The second syntactic difference distinguishing SUBJ and OBJ is at the core of our concern in this chapter. We illustrate it in (6a-c), where OBJ has an inanimate referent, in contrast to (7a-c), where the *a*-marking is traditionally attributed to semantically human/animate OBJs. Inanimate OBJs such as *coche* ('car') in (6b), on the other hand, would be ungrammatical with such mark. The completeness and coherence of (6c) is provided by the ACC clitic *lo* ('it') ensuring that the transitive V *pintó* ('painted') does have an agreeing OBJ anaphora.

- (6) a. Juana pintó un coche
 [Joan painted a car]
- b. *Juana pintó **a** un coche
 [Joan painted (*a*) a car]
- c. Juana lo pintó
 Juana it-ACC.SG.MASC painted
 [Joan painted it]

On the other hand, the human (and specific) OBJ *a su hermano* ('her brother') is correctly marked, differently from the inanimate OBJ which requires no *a* mark.⁴ In any case, it is possible to confirm that *un coche* ('a car') in (6a) and *a su hermano* ('her brother') in (7a) are both OBJ because both their equivalent sentences without the overt OBJ in (6c) and (7c) use the same ACC clitic *lo* ('it'/'him').

- (7) a. Juana pintó **a** su hermano
 [Juana painted (*a*) her brother]
- b. *Juana pintó su hermano
 [Joan painted her borteher]
- c. Juana lo pintó
 Juana he-ACC.SG.MASC painted
 [Juana painted him]

However, comparing the sentences in (7) with that in (8) we find the same *a* mark for OBL_{DAT}:

⁴ Inanimate OBJs may be marked with *a* when the speaker wishes to give them a human character, as in *era como si abrazase a un árbol o a una roca* ['it was like embracing 'a' a tree or 'a' a rock'] (Real Academia Española 2010: 659).

- (8) Juana le dió el sombrero a Pancho
 Juana he-DAT.MASC.SG gave the hat to Pancho
 [Juana gave the hat to Pancho]

Now, according to Green (2011: 213) “since *a* is also the preposition used to introduce datives there is no overt difference between the majority of direct and indirect objects.⁵ Whether the categories have genuinely fused or are merely obscured by surface syncretism is hard to say.” Green then goes on to point out that, whereas most of Latin American varieties maintain separate third person pronominal ACC and DAT clitics, much of Spain has lost this distinction. This last point finds some confirmation in Hualde, Olarrea & Escobar (2001: 342-360), who describe a range of variations in the use of ACC and DAT clitics in their overview of bilingual and contact situations involving Spanish in Latin America, as well as Central America and the USA.

To make DOM more of a puzzle, as we have already mentioned, animate OBJs are not always *a*-marked, as (7) above may imply. The OBJ there is specific enough (*Juana's brother*), but animate OBJs need not be marked if they are not specific, as (9a) shows in contrast with (9b), where the speaker has in mind specifically a lawyer or a particular kind of lawyer, even if not explicitly declared. In (9c) however we find evidence against the necessity of an animate OBJ having to be specific to be *a*-marked, as one could hardly say that the lawyer referred to is specific.

- (9) a. necesito un abogado
 [I need a lawyer] (*any, not specific, not a-marked*)
- b. necesito a un abogado
 [I need PREP a lawyer] (*specifically a lawyer, not a doctor, a-marked*)
- c. no necesito a ningún abogado
 [I don't need PREP any lawyer] (*animate, not specific, a-marked*)

Thus, neither animacy nor specificity turn out to be a failproof guide to *a*-marking of OBJ. What may then be the function of the *a*-marker in (9b-c)? It would seem that this set of contrastive examples actually supports Dalrymple & Nikolaeva's proposal that it should be treated as a marker of topicality.

Another somewhat related motivation for *a*-marking is what Tippets (2011) calls ‘relative animacy’ of SUBJ and OBJ as in (10a), where the *a*-marked OBJ *todos*

5 In the LFG framework used in this book, direct (ACC) and indirect (DAT) objects are OBJ and OBLDAT respectively.

(*everybody*) is higher in the animacy scale (human) than the SUBJ *la lluvia* (inanimate) and is made prominent by means of the prepositional marker. This could be covered under the ‘animacy’ feature of OBJ, disregarding the ‘inanimacy’ of the SUBJ. However, [-animate] OBJs may also be *a*-marked, as in (10b), where both SUBJ and OBJ are neither ‘personal’ nor animate, and neither could be said to be specific (cf. also the anthropomorphizing example in note 4).

- (10) a. *la lluvia mojó a todos*
 [the rain drenched PREP everybody]
- b. *la noche sigue al día*
 [the night follows PREP the day] (*al* = preposition *a* + article *el*)
- c. *quiero un perrito*
 [I want a puppy dog]
- d. *quiero a un perrito*
 [I love PREP a puppy dog]

This is an interesting cue to some of the language-internal reasons why *a*-marking may be used: given its flexibility, Spanish word order may be an insufficient cue to subjecthood, or objecthood, so morphological highlighting (or prominence) comes to the rescue and disambiguates for the listener which of the two participants is *not* the SUBJ – recall that SUBJ is an obligatory argument, unrestricted, and never prepositionally marked. Again, we have a case where topicality seems the neater functional explanation for *a*-marking of OBJ.

Finally, for our brief survey of *a*-marking, the lexical meaning of the V itself may also guide the choice of DOM in subtle ways, as shown in (10c-d), where the intended meaning of *quiero* (‘I want’ vs ‘I love’) guides the choice.

In sum, Spanish DOM is neither as straightforward nor as uniform as the learner or the analyst might wish. Nevertheless, the intersection of semantic (animacy) and discourse-related factors (specificity, SUBJ-OBJ relationship, secondary prominence) seem to lend support to Dalrymple & Nikolaeva’s (2011) claim that marked OBJs are associated with the information structure role of topic.

Given such complexities it is interesting to look into what the actual DOM behaviour of native speakers might be and how children learn it. Fortunately, following Laca’s (2006) lead and other researchers’ attempts to characterise factors that are relevant to Spanish DOM beyond anecdotal and constructed examples, Tippetts (2011) conducted a quantitative investigation of spoken corpora from three different major Spanish-speaking metropolitan areas: Buenos Aires, Mexico City and Madrid. He includes in his analysis only Vs that display nonuniform marking of

OBJ, and quantifies them over four basic factors: (i) animacy of OBJ, whether animate or inanimate; (ii) specificity of OBJ, whether specific or nonspecific; (iii) form of OBJ, whether a proper N or a lexical N; and (iv) relative animacy, that is, the degree of animacy of OBJ relative to SUBJ – the latter being a discourse related ‘global’ factor, compatible with Darlymple & Nicolaeva’s (2011) relating DOM to secondary topics (cf. also Leonetti 2004). Tippetts (2011) found a large measure of agreement on animacy⁶ as the primary factor favouring DOM across all three cities. He also found, significantly, that relative animacy is a stronger factor in Buenos Aires and Madrid than ‘local’ (i.e., lexical) animacy. In Tippetts’ own words, relative animacy

compares the animacy status of the subject and the animacy status of the DO using the animacy scale *human* > *animate* > *inanimate* (see Comrie 1979; Næss 2007). For example: *El partido reúne a los amigos*. In this case the DO, *amigos* (human), is higher than the S, *el partido* (inanimate), and is coded as *a* as such.” (Tippetts 2011: 109)

Furthermore this discourse-based factor, that is, relative animacy, ranks first in both Buenos Aires and Madrid, and second in Mexico City. Animacy of OBJ, on the other hand, ranks first in Mexico City and second in the other two cities. Another significant surprise in this survey is the absence of specificity of OBJ as a factor in the Madrid corpus. Finally, marked OBJs are between 33% and 39% of all OBJs across corpora (bearing in mind that Tippetts 2011 considers only Vs that do present *a*-marking across corpora). Interestingly, between a minimum of 17% and a maximum of 28% of animate OBJs are not *a*-marked across dialects. Thus Tippetts recognises the prescribed use (animate OBJs should be marked and inanimates should not), but he also laments that such prescription often fails to reflect actual usage, and summarises the situation as follows:

It is not unusual to find human accusatives unmarked as well non-human accusatives marked irrespective of animacy in spoken and written Spanish. Additional “exceptional marking” may occur as a result of ambiguities that arise from the relative flexibility of subject and object position and verbs with the *personal a* also serves a disambiguating function (2011:107).

Thus, Tippetts’ survey confirms, on the one hand, the strong presence of DOM in Spanish, but also its variation in usage across Spanish-speaking countries, with statistically significant differences between the three ‘dialects’ (Tippetts’ term) considered, as well as a clear connection with discourse.

⁶ Inanimate OBJ, albeit infrequently, is confirmed to be *a*-marked between 5% and 8% in the three corpora.

It is not surprising, then, that the conditions under which OBJ must be marked, must not be marked, or is optionally marked, have been vigorously discussed by many scholars, constituting arguably one of the most debated topics in Spanish grammar over the last 200 years, as noted by Rodríguez-Mondoñedo (2008). Recent work (e.g., Torrego 1998) attempts to tackle this issue by identifying a number of conditions that influence DOM in Spanish, such as animacy, specificity, agentivity and the semantics of the predicate, that is, the type of V involved. Each of these conditions, if valid in many or most uses of DOM, does not appear to exhaust the full range of uses (cf. Rodríguez-Mondoñedo 2008). On the other hand, Tippets' work clearly points towards an account that may not be exhaustively satisfactory by appealing simply to a number of structural conditions, however refined, without bringing to bear regional variation and discourse conditions.

We will shy away from reviewing further the huge literature on Spanish DOM here, and note, much more modestly, that its reliance on a particular lexical feature such as animacy, coupled with its relative indeterminacy – Aissen (2003) would say 'fuzziness' – and its relation to topichood, makes Spanish DOM quite complex to learn in spite of its deceptively simple morphological form. Taken together, these factors place DOM high in the Spanish acquisitional schedule because it presupposes, minimally, that functional assignment is in place. In fact, the learner must be able to distinguish not only between SUBJ and OBJ, but also between different types of OBJ (animate vs inanimate), their specificity, and ultimately their discourse status as a possible secondary topic, which we mark here as TOP₂. Furthermore, learners will need to disentangle *a*-marked ACC from the equally *a*-marked DAT (OBL_{DAT} is also typically human) in order to use the appropriate clitic anaphora – a moving target itself, at least in Spain, as Green (2011) points out.

Now let us consider briefly the acquisition of DOM by various types of learners: L1, L2 and Heritage Language speakers (HS). Rodríguez-Mondoñedo (2008) notes that, in spite of the vigorous debates on Spanish DOM, not much is known about its acquisition by children. In looking at the available CHILDES corpora, this scholar found sufficient longitudinal data in four Spanish L1 children, and counted all their instances of OBJ that should or should not be *a*-marked to see whether errors may be detected. For this reason he discounts, to our mind unfortunately, all cases where the marking is optional. From his careful analysis, Rodríguez-Mondoñedo (2008) comes to the conclusion that "children master Spanish DOM with a performance virtually errorless." This conclusion, however, is based on a conflation of results for both default and nondefault cases. Of course, both must be accounted for, but one thing is to note that children do not mark what is not marked in the input (the default case), and another is to find out how they learn to deal with the marked cases. Averaging all out, the strength of the default may well obliterate important facts. Wholly constructed from Rodríguez-

Mondoñedo's (2008) data, although following the presentation commonly used in this volume and elsewhere for SLA results, in (11) we present the pattern for his marked cases only.

(11) *DOM in L1 Spanish children (after Rodríguez-Mondoñedo 2008)*

A-MARKING OF OBJ	REQUIRED AND MARKED	REQUIRED AND NOT MARKED	NOT REQUIRED AND MARKED
MARIA	+24	-2	>2
KOKI	+12	-4	>3
JUAN	+8	-0	>2
EMILIO	+1	-1	>3

What seems to be clear from this perspective is, first, that the acquisitional pattern is neither as strong nor as uniform across these four children as the author asserts, and secondly, that the error rate for some of them is not insignificant, particularly for Emilio and Koki. Emilio, as it turns out, marks DOM at less than chance. Why? Rodríguez-Mondoñedo (2008) explains that Emilio is a Spanish speaking boy living in a Catalan community, and that Catalan does not have DOM, then he adds: "It has been observed already that Spanish-speaking children in bilingual communities (when one of the languages does not have DOM) drop *a* occasionally (Silva-Corvalán 1994; Luján 1996; Montrul 2004 ...)". The other interesting case is Koki, a Mexican child with Spanish-speaking mother and American English-speaking father. Previous to the recordings in Mexico, the family lived for the first six months in Poland, the next six in Argentina and two months in the US. Koki's accurate use of DOM is lower than that of the other children, except Emilio. The bilingual home and the different environments of exposure may have contributed to her pattern of DOM acquisition. This is an interesting set of facts, which together with Emilio's may help understand the consistently low accuracy with DOM – compared to native speakers within a single environment – that Montrul and her colleagues (e.g., Montrul 2008; Montrul & Bowles 2009; Montrul & Sanchez-Walker 2013), after Silva-Corvalán (1994), repeatedly find in the US with young HS, whether simultaneous or sequential bilinguals, including adult immigrants after a certain period of residence in the US.

Montrul and her colleagues attribute the difference in DOM to 'incomplete acquisition' or later 'attrition' (in the case of adult immigrants), which in turn may be due to the reduced quantity and quality of the input received by the HS bilinguals from adult HS whose spoken Spanish may itself be 'attrited' (except for recent adult arrivals). Young HS bilinguals may also undergo attrition, not unlike adult HS, under the influence of, and in convergence with, English, the dominant lan-

guage, which has no DOM. But, as Montrul & Sanchez-Walker (2013: 127) candidly admit, DOM in Spanish “is conditioned by unclear semantic and pragmatic factors. We still do not know whether DOM is primarily regulated by type of verb, type of object, other elements in the sentence or all of these”.

The nature of the Spanish-speaking environment should also be taken into account, we believe. One would wish to ask which speakers do HS communicate with in their Heritage Language? In the US there are Spanish speakers from all over the Spanish-speaking world, which would surely not be a homogenous Spanish-speaking community, as Tippets (2011) demonstrates specifically regarding the way DOM is used. On the other hand, Mexican native speakers living in Mexico can be assumed, unlike their US counterparts, to be using their language in a homogenous community, so a comparison between the two groups of speakers needs to be moderated by this basic difference. Overall, then, it would be surprising if HS, adult or younger bilinguals, in the US used DOM according to the structural factors (such as animacy and specificity) indicated by Montrul and her colleagues. Variation, as found in their results, seems to be a reasonable expectation. Furthermore, the type of task (grammaticality judgement) and the presentation of results by group means rather than by individual speakers used in their research paradigm would tend to obscure critical developmental or environmental (sociolinguistic) differences among speakers, resulting from their individual experience and pattern of exposure, as we just saw with the four children looked at by Rodriguez-Mondoñedo (2008).

Naturally enough those who have worked with Spanish L2 have found DOM rather elusive and difficult for the learners. We will look here at two studies, that is, one by Guijarro-Fuentes (2011, 2012) and the other by Farley & McCollam (2004). Guijarro-Fuentes proposes an interesting approach which he calls the Feature (In)accessibility Hypothesis, whereby DOM is the result of a number of interacting features subsumed by OBJ, some of which are structurally given (such as the grammatical relationship with the V, e.g., its ‘direct’ rather than ‘oblique’ nature of the relationship), and others are inherent in the particular OBJ (such as animacy and specificity). The latter features, which Guijarro-Fuentes calls ‘interpretable’ features in his research paradigm (cf. Lardiere 2008), are not all equally accessible to learners. He finds that ‘animacy’ is the most accessible (or learnable) of these interpretable (or inherent) features.

Farley & McCollam (2004) report on an experiment involving 29 students. They explicitly use DOM (which they call ‘personal *a*’) and subjunctive marking to test PT’s Teachability Hypothesis (Pienemann 1984). Their conclusion is that results do not support PT as presented by Pienemann (1998) and call into question the hierarchy for Spanish proposed by Johnston (1995). Following Johnston’s developmental hypotheses, Farley & McCollam reason that ‘personal *a*’ is at a much lower stage (Johnston’s stage 4, corresponding to the phrasal pro-

cedure stage in this volume) than subjunctive marking (Johnston’s stage 7, corresponding to the S-BAR procedure stage). Then, following the logic of Pienemann’s (1984) Teachability Hypothesis (i.e., learners cannot ‘skip’ a stage even with instruction), they ensure that their learners are all ‘ready’ to learn Johnston’s stage 4 (personal *a*), but that, conversely, none is ready for subjunctive marking on V in subordinate clauses. The students are then divided into four groups of 6 or 7 students each: three of the groups were instructed in each of the two structures with different treatments (explicit instruction, structured input, and processing instruction, VanPatten 1996, 2007), and the control group received no instruction on those structures. Results are collated in (12), which shows the number of learners who are deemed developmentally ready for the relevant structure, and the number of learners who have actually developed it after the treatment. Results are clearly more favourable for subjunctive marking (c. 60%) than for ‘personal *a*’ (c. 40%), even though the learners were deemed to be ready for DOM but not for subjunctive.

(12) *Global results from Farley & McCollam (2004) teachability experiment*

STRUCTURE	TREATMENT & LEARNERS	READINESS	DEVELOPMENT
OBJ MARKER	INSTRUCTED (N. = 6+7+6)	19	7
	CONTROL GROUP (N. = 6)	6	0
SUBJUNCTIVE MARKER	INSTRUCTED (N. = 6+7+6)	0	11
	CONTROL GROUP (N. = 6)	0	0

Given our discussion above regarding the nature of DOM, it seems that Farley & McCollam (2004) place it too early in the developmental path. DOM is not just morphological marking of OBJ when it is animate because, as pointed out above, not all animate OBJs are *a*-marked – they must also be specific, among other things. Furthermore, some of the tasks used in their study require a much higher structure involving the primary topicalisation of *a*-marked OBJ plus the clitic_{OBJ}, as in (13), which introduces an extra complicating factor.

- (13) a los viejos no los respetan los jóvenes
 a-DOM the old-PL.MASC no they-CL.ACC.PL.MASC respect-3.PL the young-PL.MASC
 [(as for) the older people, young people don't respect them]

This is definitely not a phrasal procedure stage structure because it calls for higher level resources, in particular the S-procedure and full functional assignment (cf. Di Biase & Kawaguchi 2002, and § 3.1, ch. 3 this volume, for a parallel structure in

Italian). That is, the learner must distinguish not only SUBJ from OBJ, but also differentiate among types of OBJ, one of which must be also distinguished from OBL_{DAT}. The initial XP *a los viejos* ('[a] the elderly') looks like a 'garden path' DAT. If differentiation fails, the clitic may be wrongly case-assigned. On top of all that, the construction in (13) also displays marked word order: TOP_{OBJ} Cl_{ACC}-V SUBJ. Hence, it is definitely on the higher stages of L2 learning. This discussion naturally leads into our PT-based hypothesis, which takes into account the relevant facts for a more precise location of DOM in the development of L2 Spanish.

3. The developmental hypotheses

The structure we are focusing on in this chapter – that is, the language-specific DOM – is expressed by means of a morphological marker on OBJ. As such, we begin by placing it in (14) within our morphological development hypothesis for Spanish L2, in line with the current proposals for PT outlined in chapter 1 of this volume (cf. also the Italian schedule, ch. 3). That is, the morphological schedule, whose structures are obligatory, is presented separately from the syntactic schedules based on the Prominence Hypothesis and the Lexical Mapping Hypothesis which, on the other hand, represent speakers' choices largely related to discourse-pragmatic factors. Hence, given that DOM sits at the intersection of the syntax-semantics and syntax-discourse interfaces, we need to discuss it within the Prominence Hypothesis and the Lexical Mapping Hypothesis.

Like other morphological schedules in previous chapters of this volume, the one for Spanish exhibits a developmental path beginning with single words and formulas followed by lexical-level morphemes such as the plural *-s* in Ns. Spanish has a largely regular masculine/feminine gender marking and its plural morpheme *-s* is fairly easy to segment. For example, *maestra/maestro*, female and male 'teacher' respectively, become *maestras/maestros* for their plurals. For these procedures requiring lemma access and category marking no unification of features is required.

Learners who are able to process the next procedure, that is unification of features within the phrase, are minimally able to unify the value of one feature, such as [plural], at the phrasal node. Our schedule in (14) reflects this agreement at the NP node where the attributive plural form *americanos* ('American') agrees with the head N *coches* ('cars'), or the VP agreement between the predicative adjective *grandes* ('big' PL) with the plural form of the copula *son* ('are'). This is the stage at which Farley & McCollam's (2004), following Johnston (1995), placed DOM. In our view, it is possible that *a*-marked OBJs may appear at this stage – as well as OBLs expressed by PPs (e.g., dative, locative) – but only if DOM were understood simply as a structural morphological mark-

(14) Developmental stages hypothesised for L2 Spanish morphology, identifying the location of DOM

PROCESSING PROCEDURE	MORPHOLOGY OUTCOME/ STAGE	SPANISH-SPECIFIC STRUCTURE	EXAMPLE
S-BAR PROCEDURE	INTERCLAUSAL MORPHOLOGY	subjunctive marking in subordinate clause	<i>me parece ridículo que cada familia tenga dos coches</i> [to me it seems ridiculous that each family has two cars]
S-PROCEDURE	INTERPHRASAL MORPHOLOGY	PP _{TOP} (SUBJ) Cl _{ACC} V [= DOM]	<i>a la maestra Juan la vió ayer</i> [(PREP) the teacher, Juan saw (her) yesterday]
		XP _{TOP} (SUBJ) Cl _{ACC/DAT} V	<i>Salzburg lo he visitado una vez</i> [Salzburg I have visited once]
		SUBJ _{PL} COP _{PL} PRED ADJ _{PL}	<i>los perros son buenos</i> [the dogs are good]
		SUBJ _{PL} V _{PL} agreement	<i>los niños corren</i> [the children run 3 PL]
PHRASAL PROCEDURE VP	PHRASAL MORPHOLOGY	COP _{PL} ADJECTIVE _{PL} agreement	<i>son grandes</i> [are big PL]
PHRASAL PROCEDURE NP		NOUN _{PL} ADJECTIVE _{PL} agreement	<i>coches americanos</i> [American PL cars]
CATEGORY PROCEDURE	LEXICAL MORPHOLOGY	plural marking on nominals	<i>árboles</i> [trees]
			<i>bonitos</i> [beautiful PL]
LEMMA ACCESS	INVARIANT FORMS	single words, formulas	<i>hola</i> [hallo] <i>me llamo Pat</i> [my name is Pat]

er within the VP. However, we have seen that other syntactic considerations are involved, which require full functional assignment, e.g., the differentiation of OBJ from SUBJ via SUBJ-V agreement, which in turn relies on the S-procedure being in place. This is where the positional lock on SUBJ and OBJ is opened and even a postverbal SUBJ may be still identified as such. Thus DOM is best placed at the S-procedure stage, where unification happens at the node(s) adjoining different phrases.

In our schedule then, at the S-procedure stage, as well as sentences exhibiting the familiar SUBJ-V agreement, we may find also DOM, both in its canonical postverbal position (but note that OBJ does not require any morphological

agreement in such case) or as primary TOP in preverbal position, in a sentence equivalent in structure to (13), an example borrowed from Farley & McCollam's (2004) DOM tasks. In this case, unification must occur between the gender and number of the N in the NP *a la maestra* ('PREP the teacher') and the ACC clitic of V. The full justification for placing DOM here rather than at the phrasal procedure stage will be clearer shortly when we deal with syntax.

The final stage on our morphological schedule is aligned with Farley & McCollam (2004) and Bonilla (2015), and concerns the S-BAR procedure stage, which in Spanish contains those constructions in which the relationship between main and subordinate clauses must be marked through V inflection. Thus, in (14), the subjunctive V form *tenga* ('should have') in the subordinate clause is compatible with the epistemic *me parece ridículo* ('to me, it seems ridiculous') in the main clause.

As we argued above, DOM is an optional structure sitting at the interface of the syntax-discourse and syntax-semantics interface, hence it will be further discussed in the next two schedules for syntactic development. In (15) the developmental progression covered by the Prominence Hypothesis for declarative sentences is broadly laid out (cf. § 4.2, ch. 1, this volume, and also §§ 3.1, ch. 2 for English, ch. 3 for Italian, and ch. 4 for Japanese).⁷ This hypothesis relates to word order, canonical and otherwise, which respond to discourse-pragmatics constraints. These may also operate through other means, and prominence may be achieved by means of prosody, for instance, or indeed morphology, as well as word order. Thus we now focus on the location of DOM within the Prominence Hypothesis schedule in (15).

In this schedule we hypothesize, as with other languages, that learners will start from the canonical word order, which in Spanish is SVO. Furthermore, because Spanish is a prodrop language, the majority of canonical utterances are VO (cf. Bonilla 2015). In principle there is no reason why DOM should not be located here. However, at this stage learners have not learned yet to differentiate anything much inside the canonical order, and distinguish OBJ from SUBJ merely by their position. So all OBJs are equally underspecified and their animacy or inanimacy would make no difference for early-to-high intermediate learners, especially if their L1 does not differentiate such OBJs (which is the case, e.g., for English, Japanese and Catalan). Likewise, at the next stage, a topicalised adverbial phrase leaves the core elements of the canonical block unaltered. So, if a potentially differentiable animate OBJ is produced, again the learner would not

7 Questions are extensively treated elsewhere in this book; cf. ch. 2 for English, and ch. 8 for Italian respectively, but also ch. 9 for German.

(15) *Developmental stages hypothesised for L2 Spanish syntax, identifying the location of DOM within the Prominence Hypothesis*

STAGE	STRUCTURE	SPANISH EXAMPLE
NONCANONICAL WORD ORDER	(PP _{TOP}) (SUBJ) Cl _{ACC} V ADJ = [DOM]	<i>a la maestra Juan la vio ayer</i> [(PREP)the teacher Juan saw (her) yesterday]
	SUBJ V OBJ _{TOP2} = [DOM]	<i>Juan vio a la maestra</i> [Juan saw (PREP) the teacher]
	XP _{TOP} (SUBJ) Cl _{ACC/DAT} V	<i>Salzburg lo he visitado</i> [Salzburg I have visited]
XP _{ADJ} CANONICAL WORD ORDER	XP _{TOP} (SUBJ) V OBJ	<i>ahora Juan tiene dos hermanos</i> [now Juan has two brothers]
		<i>en Sydney tengo bicicleta</i> [in Sydney I have a bicycle]
CANONICAL WORD ORDER	(SUBJ) V OBJ	<i>Juan tiene amigos</i> [Juan has friends] <i>Juan tiene coche</i> [Juan has a car]
LEMMA ACCESS	<i>words, formulas</i>	<i>me llamo Pat</i> [my name is Pat]

be able to differentiate it inside the blocked string. It is important to note, however, that at both stages we will find animate arguments which are not *a*-marked but nevertheless grammatical, such as in (16).

- (16) a. Juan tiene amigos
[John has friends]
- b. ahora Juan tiene dos hermanos
[now John has two brothers]

As we have seen in § 2, the feature [+animate], however central, is insufficient by itself for OBJ to be *a*-marked and needs to be accompanied by other features such as [+specific], which is linked to discourse. So we place DOM at the next stage, the noncanonical word order stage, even though its descriptor may not be quite adequate given that the actual word order for sentences like *Juan vio a la maestra* ('Juan saw PREP the teacher') looks perfectly canonical. Recall that passives also look pretty canonical in this regard (e.g., in English as well as in Spanish), but this does not mean that we place them at the default stage. In any case, this is the stage at which full (i.e., independent of position) functional assignment must be assumed to free up word order, because GFs are marked morphologically or by other means. The learner at this stage must also be able

to differentiate core arguments from each other and from non-core arguments as well as match prominence features, that is, [\pm prominent, \pm new], with the appropriate constituent, whichever the order. Notice in particular that this *a*-marking has little or no role to play in the grammatical relations of the OBJ (which remains OBJ whether or not it is *a*-marked). What the preposition marks here, by morphological means, is the 'secondary prominence' of an argument mapped onto an OBJ that shares in the 'aboutness' of the sentence, and often, but not always, is animate and agentive-like (cf. morphologically marked prominence in Japanese, ch. 4, this volume). This *a*-marking is maintained also when the speaker promotes such a differentially marked OBJ as primary topic, as in (13). So our representation of constructions such as these shows up in both the morphological and the syntactic schedules, because the former, in (14), reflects the required feature unification ensuring that the TOP constituent is correctly interpreted as NON-SUBJ, and the latter, in (15), reflects discourse-generated word order differences. In the end, because we characterise the differentiated OBJ as bearing a TOP₂ function in the spirit of Dalrymple & Nikolaeva (2011), it seems reasonable to place it within the Prominence Hypothesis schedule and at the noncanonical stage, in contrast to earlier interpretations of it as an exclusively morphological structure.

Let us now turn, briefly, to the Lexical Mapping Hypothesis for Spanish in (17), because DOM straddles across the syntax-semantic interface as well. We noted for instance that the largest measure of agreement on the actual usage of DOM reported by Tippetts (2011) across three metropolitan varieties of Spanish relates to a lexical semantic feature, that is, [animacy], with the highest ranking for the discourse-based 'relative animacy' in two of the metropolitan areas considered (Buenos Aires and Madrid) and second rank in the remaining area (Mexico City).

After the universal single word and formula stage, the development of lexical mapping begins with V representing actions/states and with default mapping of the higher ranking agent/actor on N preceding the theme/patient represented on the other N. The prevailing VO default is unlikely to pick a differentially marked OBJ since the majority (61% to 67%, according to Tippetts 2011) of OBJs in the input are not marked. Hence learners are hypothesized to treat all OBJs in the same way, whether animate or inanimate, specific or not specific, definite or indefinite. In other words, the links to critical semantic features and discourse are, at this stage, absent. Proceeding upward, learners begin to increase their stock of complements with arguments additional to their default string. Crucially, additional arguments must be differentiated from the default core arguments, particularly from OBJ given the prodrop nature of the Spanish SUBJ. This is exactly the stage where the semantics of the additional arguments, through their prepositions, makes more transparent their relationship with V. So OBL_{DAT} for goal or beneficiary will com-

(17) *Developmental stages hypothesised for L2 Spanish syntax, identifying the location of DOM within the Lexical Mapping Hypothesis*

STAGE	CONSTRUCTIONS	SPANISH EXAMPLE
NONDEFAULT MAPPING	unaccusatives, passives, causatives, impersonals, exceptional V constructions	<i>Juan fue empujado por el ladrón</i> [Juan was pushed by the robber] <i>a ella le gustan las plantas</i> [she likes plants] <i>disminuye el tráfico</i> [traffic decreases]
	DOM	<i>yo conozco a un médico</i> [I know (PREP) a doctor]
DEFAULT MAPPING AND ADDITIONAL ARGUMENTS	agent/experiencer mapped on SUBJ, patient/theme on OBJ and other roles on OBL arguments	<i>José da pan a los niños</i> [José gives bread to the children] <i>ahora vivo en Barcelona</i> [now I live in Barcelona]
DEFAULT MAPPING	agent/experiencer mapped on SUBJ, patient/theme mapped on OBJ	<i>Juan tiene amigos</i> [Juan has friends] <i>Juan compra manzanas</i> [Juan buys apples]
LEMMA ACCESS	single words, formula	<i>me llamo Pat</i> [my name is Pat]

monly be animate, as in (18a), whereas OBL for location, instrument, and so on, will not, as in (18b).

- (18) a. José da pan a los niños
[José gives bread to the children]
- b. ahora vivo en Barcelona
[now I live in Barcelona]

These OBLs are usually processed after the core argument(s) and are differentiated from OBJ on account of position, as well as by the presence of a preposition. Given that this stage is still characterised by default mapping (plus additional argument), it is unlikely that DOM could be acquired at this point, because the semantic role mapped on OBJ (theme or patient) is still assigned by default and the additional arguments need to be differentiated from it. Nevertheless, the marking of different OBLs with a variety of prepositions is another step towards distinguishing animate arguments (other than SUBJ) from inanimate ones. In other words, the learner is moving towards building the semantic interface with syntax, a resource for becoming able to mark OBJs differentially.

The next and highest stage on this schedule includes all nondefault mappings which account for alternative, pragmatically motivated constructions such as passives and complex constructions such as causatives. This higher stage is where hierarchical correspondences (such as the agent role mapping on the SUBJ function) are subverted if the speaker's communicative intention or lexical choice requires it. Hence we can expect at this stage to find scope for a more confident differentiation of arguments, including differentiation between OBJs, mostly involving animacy and specificity among other constraints. It seems to us then a fairly safe stage at which to place the emergence of DOM on this developmental schedule, although its marginal optionality may still elude or confuse the learner for a long time, leading to persistent errors.

In sum, our presentation of the three developmental schedules for morphological development and syntactic development along the Prominence Hypothesis and the Lexical Mapping Hypothesis allows us to hypothesize, in (19), an implicational hierarchy for acquiring DOM in L2 Spanish.

(19) no marking of OBJ > marking of [\pm animate] OBLs > DOM > topicalised DOM

Overall one must bear in mind, however, that achieving either the S-procedure stage or the distinction between animate and inanimate arguments does not guarantee that DOM will be fully deployed by the learner, given, among other things, its interface with discourse prominence. Positive evidence will be the morphological *a*-marking of OBJ with transitive Vs. A higher level of evidence is obtained when the learner also produces a TOP which is *a*-marked (even if inanimate), which is coreferential (i.e., it shares number and gender features) with the ACC clitic marker of OBJ on the V, as in (13) above.

4. The study

Our cross-sectional study of six learners of Spanish L2 is guided by a single research question: do learners of Spanish L2 acquire DOM at the hypothesised developmental points in the PT schedules proposed in § 3?

4.1. Informants and tasks

The six learners in this study are Austrian German-speaking students. Their participation in the research project was voluntary and they all signed the relevant consent form. All of them are 19-year old females who have been studying Spanish in a formal school setting for three years, totalling approximately 260 hours of

instruction, attending the same Spanish class in the same school given by the same teacher and thus having been exposed to the same classroom material. In addition to formal instruction, half of the students had gained exposure to natural input in Spain: two of them, BEN and JAN (all names are code names) spent three months in Spain, and one (VES) spent one month there, as can be gleaned from a questionnaire administered to participants to gather information about their language background. Instruction sessions over the two semesters of observation were digitally audio-recorded and later transcribed and analysed. From this analysis it transpires that explicit instruction on the subjunctive and DOM, the structures on focus in this article, did not receive the same time allocation. Subjunctive constructions were dealt with extensively from the end of the second year of instruction onwards, whereas DOM was only dealt with in the second year and to a much lesser extent (Hinger 2011).

In order to elicit spontaneous oral production data, the learners were administered two tasks in a single session outside of their classroom environment for 15 minutes on average. These task sessions also were digitally audio-recorded and later transcribed and analysed. The first task is a classical communicative information gap taken in pairs, the second a monologic task. Students picked their pair themselves with no linguistic criteria applied in reference to pairing. The two tasks consist of written and visual prompts “providing information about the context, the content and the purpose” (Tankó 2005: 42) of the oral production. The topics chosen referred to contents of the ongoing Spanish classroom teaching, though neither the task types nor the task goals and requirements had in any way been rehearsed. The monologic task aims at eliciting oral production relating to the environmental problems caused by pollution, and the learners are given a visual and written prompt (in Spanish) on the consequences of pollution for humanity, and are encouraged to provide ideas to help stopping climate change. The interactive task focuses on the situation of renting a room in the target country, namely in the capital of Spain, Madrid. The two partners were given different roles described on cards written in Spanish. Partner A was provided with the information of living on her own in a flat which she wants to share with a Spanish native speaker because she spends some time in Madrid in order to improve her language knowledge. A picture of the flat was also shown to partner A for her to describe the flat to partner B, who phoned after reading an announcement in the paper. So, partner B has the role of a Spanish native speaker hunting for a room in Madrid. She is especially attracted by an announcement in the paper made by an Austrian because she had spent some time in Austria and was keen to practice her German. In the end, she decides not to hire the room and needs to find some explanations. In both tasks, the researcher role was that of facilitator and occasional interlocutor in order to help overcome situational problems such as anxiety by reassuring and motivating learners, nonverbally and verbally, to continue their talk.

The tasks were successful in eliciting a corpus of a total of 4,564 tokens and 1,529 types of Spanish lexical items distributed in 486 clauses, an average of about 94 clauses per informant, as shown in (20). The minimum number of clauses uttered is 35 by GRA, the maximum 151 by JAN. These differences cannot be attributed to the three month experience in Spain by JAN because informant BEN, for instance, also spent the same amount of time in the target country and produced 80 clauses and 738 tokens, both being less than those of informant ROS, who had never visited a Spanish-speaking country.

(20) *The corpus for the six learners*

	GRA	VES	TAN	ROS	BEN	JAN	TOTAL	MEAN
WORD TOKENS	370	611	376	1.025	738	1.444	4.564	760.67
WORD TYPES	164	224	174	323	241	403	1.529	254.83
CLAUSES	35	55	49	116	80	151	486	93.90

4.2. Results and analysis

In this section we present the results concerning first morphological development and then syntactic development.

With regard to their current morphological development in terms of classic PT (Pienemann 1998), all of the informants have achieved the S-procedure stage, and three of them also the S-BAR procedure stage, as shown in (21).

(21) *Cross-sectional study of morphological development in L2 Spanish*

STAGE	STRUCTURE	GRA	VES	TAN	ROS	BEN	JAN
S-BAR PROCEDURE	subjunctive marking in subordinate clause	-2		-4	+3 -3	+2 -1	+4
S-PROCEDURE	TOP _{PP} Cl _{ACC} V (SUBJ)						
	TOP _{XP} Cl _{ACC/OBL} V (SUBJ)			+2	+1		+3
	SUBJ _{PL} COP _{PL} adjective _{PL}	+1 -2	+1 -1	+1	+2 -2	+1	+4
	SUBJ _{PL} V _{PL}	+2	+8	+4	+6	+3 -1	+12
PHRASAL PROCEDURE	COP _{PL} adjective _{PL}	+1	+2	+1		+2	+3 -1
	N _{PL} adjective _{PL}	+5	+11	+6	+9 -2	+10	+29
CATEGORY PROCEDURE	N _{PL}	+5	+5	+4	+18	+6	+20

+ = supplied; - = not supplied; empty cell = no context

Looking at the figures within the S-procedure stage, we can see that three learners (GRA, VES and ROS) are still uncertain with the agreement between SUBJ and adjective, but they do, like all the others, manage unification on the plural agreement between SUBJ and lexical V without fail.⁸ It may be the case that the copular clause absorbs greater processing resources because COP itself, placed between the SUBJ and the predicative, also has to unify its [number] feature with both the SUBJ and the predicative adjective. A possible case of a 'soft barrier' (cf. § 5, ch.1 and further ch. 2, § 2 for English, and ch. 3, § 2.2 for Italian, this volume). Nevertheless, despite the low return on this structure, some good indicators of functional assignment are in place, as exemplified in (22a-b).

- (22) a. TAN los profesores no son idoles [target: idolos]
 the professors-PL no are-PL idols-PL
 [the teachers are not idols]
- b. VES los jóvenes pueden cambiar algo
 the young-PL can-PL change something
 [young people can change something]

What seems to be rather harder to produce at this stage is topicalisation involving unification between an XP_{TOP} with the clitic marking the OBJ (23a) or OBL (23b) function on V, because the clitic must have both the grammatical function (ACC or DAT respectively) that the displaced XP would have in its default position in the sentence, and the same number and gender features as the N in the XP.

- (23) a. TAN Schwaz ah lo conozco
 Schwaz-SG.MASC ah it-CL.ACC.SG.MASC know-1.SG
 [Schwaz (a town in Austria), I know]
- b. JAN a ella también le gustan las plantas
 to her-3.SG.FEM also she-CL.DAT.3.SG.FEM like-3.PL the plants-PL
 [she also likes plants]

Topicalisations involving OBJ or OBL are quite rare: only 6 were found in the whole database, and notice also that none of the learners produced a topicalised differentially marked OBJ. These constructions are other prime candidates for 'soft

⁸ Note that SUBJ & V agreement is only counted when SUBJ is expressed lexically or pronominally. Getting the right V-form with null SUBJ is part of categorial learning (cf. § 2.1, ch. 3 on Italian, this book).

barriers' because, additionally to feature unification, they involve word order choices triggered by discourse-pragmatic constraints.

At the next, S-BAR procedure stage, we find three learners. One is ROS, who satisfies the emergence criterion for subjunctive morphology because she has three positive occurrences with lexical and structural variation (cf. § 5, ch. 1, this volume). Her accuracy, however, does not go beyond 50%. The one thing she misses in three of her otherwise well constructed sentences is precisely the morphological shape of V in the subordinate clause, as in (24), where she provides the default indicative rather than the subjunctive V-form.

- (24) ROS no creo que necesito lámparas todo el día
 no believe-1.SG that need-1.SG.INDICATIVE lights all day
 [I don't believe that I need lights all day]

The other two learners also produce several contexts of this higher level structure, and the task canvassing opinions about a controversial issue which young people care about was clearly successful in pushing informants to attempt it. However, again, like ROS in (24), BEN's morphological V-form is the default indicative in one case and the required subjunctive mood in two cases, as in (25a), whereas JAN consistently marks V in the subordinate clause with the subjunctive mood, as in (25b).

- (25) a. BEN no hace falta que cocinen una cosa exótica
 not has need that cook-3.SG. SUBJUNC a thing exotic
 [there is no need (that she) cooks an exotic thing]
- b. JAN es muy importante que la gente tenga más respecto
 is very important that the people-SG have-3.SG.SUBJUNC more respect
 [it is very important that people show more respect]

Let us now turn, to our learners' syntactic development with results shown in (26) focusing on DOM, that is, to the TOP₂ features of some animate OBJs, as identified in the Prominence Hypothesis in (15).

As we have mentioned, the overall frequencies of DOM structures are not abundant in our corpus. This is due, at least in part, to the fact that, as in other chapters of this volume, the counting for valid syntactic structures includes lexical V to the exclusion of copular and presentative sentences, whose word order responds to specific constraints and are preferably analysed separately.

All our learners produce canonical word order, which is the foundational syntactic stage, and at least one structure from the next stage, the XP canonical word order stage. Four of them also produce at least one OBJ or OBL topicalisation, as

(26) *Cross-sectional study of syntactic development in L2 Spanish identifying the position of DOM within the Prominence Hypothesis*

STAGE	STRUCTURE	GRA	VES	TAN	ROS	BEN	JAN
NONCANONICAL WORD ORDER	(PP _{TOP}) Cl _{ACC} V (SUBJ)						
	(SUBJ) V OBJ _{TOP2} [= DOM]	-1	-1	1	-2	-2	3
	(NP _{TOP}) Cl _{ACC} V (SUBJ)			2			
	(PP _{TOP}) Cl _{DAT} V (SUBJ)				1	1	5
XP _{ADJ} CANONICAL WORD ORDER	XP _{ADJ} (SUBJ) V OBJ	1	2	2	2	4	5
CANONICAL WORD ORDER	(SUBJ) V OBJ	6	6	6	11	16	15

exemplified in (23), which characterises them as higher intermediate learners. As for DOM, all learners produce at least one potential context for it, that is, a sentence with an animate OBJ with a sufficient specificity load, but not all of them *a*-mark it. GRA and VES for example never do – regardless of animacy or specificity they seem to regard all OBJs the same, as exemplified in (27a-d).

- (27) a.

GRA

tengo

otras ofertas

have-1.SG other offers (not animate)

[I have other offers]
- b.

GRA

tengo

una amiga

have-1.SG a friend (animate, not specific)

[I have a friend]
- c.

GRA

voy

a llamar

mi amiga

go-1.SG to call ø-DOM my friend (animate and specific)

[I am going to call my friend]
- d.

VES

pero ahora no tengo

otra persona

but now no have-1.SG ø-DOM other person (animate and specific)

[but now I don't have anyone else]

It is crucial to note here that (27a-d) provide a contrastive key to our analysis: as against (27a) with a [–animate] OBJ, both (27b and c) have animate OBJs. However, *una amiga* (‘a friend’) in (27b) is not specific, and is therefore counted among the 6 canonical word order structures produced by GRA in (26). By con-

trast, *mi amiga* ('my friend') in (27c) by virtue of the possessive determiner is quite specific, and is therefore counted as -1 for OBJ_{TOP2} [= DOM] because it should have been *a*-marked. Similarly, *otra persona* ('another person') in (27d) is counted as -1 because it appears to have sufficient specificity in the discourse to be *a*-marked as OBJ_{TOP2}, but VES does not *a*-mark it. An interesting case is provided by ROS in (28), where *la tierra* ('the earth') is a [-animate] OBJ, but the intended meaning ('I love the earth' rather than 'I want the earth') would likely trigger DOM (cf. the contrast in (10c-d) and note 4). Given that this learner has achieved the top stage of morphological development (subjunctive marking), it is plausible to assume that she could mark DOM when discourse conditions may require it, but she does not.

- (28) ROS yo quiero la tierra
 I love-1.SG ?ø-DOM the earth (not animate, specific, prominent)
 [I love the earth]

Four out of our 6 learners behave within this pattern of not differentiating OBJs. The exceptions are TAN and JAN, with one and three differentially marked animate OBJs respectively, as in (29a-c), where DOM is produced in what appears to be the appropriate environment for it.

- (29) a. TAN conozco a algunas personas
 know-1.SG a-DOM some people
 [I know certain people]
- b. JAN yo conozco a un médico
 I met-1.SG a-DOM a doctor
 [I met many people from all over the country]
- c. JAN conocí a mucha gente de todo el país
 met-1.SG a-DOM many people from all the country
 [I met many people from all over the country]

Notice that the two learners who produce DOM are the only ones who produce more than one topicalised OBJ (TAN) or OBL (JAN). Notice also that JAN, the only informant with sufficient evidence for marking TOP₂, does not produce a DOM structure as primary TOP, and neither does anybody else. Such production would be the best evidence that the learner may not be conflating DOM with DAT because the clitic anaphora in such cases would have to be ACC even though the coreferential DOM TOP is expressed by a PP. We nevertheless choose to leave the structure in our implicational hierarchy, to be tested on a richer database.

Out of the four learners who do not produce DOM, two of them (ROS and BEN) appear to be in a state of readiness for DOM, both producing at least one topicalised structure with a DAT clitic, as in (30a-b). Though both expressions sound fairly idiomatic, not all learners use them.

- (30) a. ROS

a ti

te

interesa

mucho

to you

you-Cl.DAT.2.SG

interest-3.SG

much

[to you it is very interesting]
- b. BEN

a mí

me

da

igual

to me

I-Cl.DAT.1.SG

give-3.SG

same

[to me it is the same]

To round up the result section we now present in (31) the analysis for the Lexical Mapping Hypothesis as shown in (17).

(31) *Cross-sectional study of syntactic development in L2 Spanish identifying the position of DOM within the Lexical Mapping Hypothesis*

STAGE	CONSTRUCTIONS	GRA	VES	TAN	ROS	BEN	JAN
NONDEFAULT MAPPING	passives and exceptional Vs		2		1	3	5
	DOM	-1	-1	1	-2	-2	3
DEFAULT MAPPING AND ADDITIONAL ARGUMENTS	agent/experiencer mapped on SUBJ, patient/theme on OBJ and other roles on OBL arguments	2	7	3	6	7	10
DEFAULT MAPPING	agent/experiencer mapped on SUBJ, patient/theme mapped on OBJ	5	8	7	11	15	15

We cannot help but notice again the paucity of production in the nondefault mapping stage. Given the highly marked or inherently lexical nature of constructions such as passives, causatives, impersonals, and exceptional Vs, this is not surprising in spontaneous speech. As a matter of fact, only JAN produces a passive construction, and all other figures in the top row of (31) refer to three Vs: *gustar* (‘like’), *interesar* (‘interest’) and *parecer* (‘seem’), which map the lower ranking theme role on postverbal SUBJ. DOM constructions are also placed in this schedule because of their semantics/syntax interface character, as discussed at the bottom of § 3. Their figures, naturally, are the same as in (24), which looks at them in terms of TOP₂. Here it is worth noting that JAN, the strongest DOM performer, is also the one with the greater range of nondefault constructions.

4.3. Discussion

Our developmental results for morphology in (21) and syntax in (26) and (31) are compatible with the PT hypotheses in (14)-(15) and, essentially, also (17). We can see from (21) that morphological resources are available to all learners up to the S-procedure stage, so they can all handle functional assignment unequivocally. Three learners have also achieved the higher S-BAR procedure stage, which means they can handle at least some subjunctive constructions. This state of affairs appears to be corroborated for syntax in (26), in so far as four learners are able to use some marked word orders, which also assume unequivocal functional assignment, and in (31), which shows they can handle nondefault mapping operations.

Despite the high stages achieved, only two of the learners apply DOM. The other four do produce some contexts for DOM, but do not *a*-mark OBJ. In both the learners who use DOM, the particular feature that seems to be relevant is animacy, but also specificity. This result partly coincides with Guijarro-Fuentes' (2011, 2012), as well as with Tippetts' (2011) indication that animacy is the primary factor favouring DOM in his corpus-based study. The other features, such as specificity, may be tied up more with the current discourse model of the speaker. So it seems appropriate to place DOM high in the syntactic sequence within the Prominence Hypothesis space, where full functional assignment is assumed, and learners become capable of assigning prominence to any arguments. We have argued that DOM is to be placed also at the top stage of the Lexical Mapping Hypothesis, and this also has some support in the results. If we treated DOM as a purely morphological marker we would place it, with Johnston (1995) and Farley & McCollam (2004), at the phrasal stage because it is part of PP within VP. As we have seen this is not the whole story, and our results confirm the multifaceted nature of DOM. At bottom its use or nonuse is at a crossroad, an interface constrained by information which is semantic (animacy), syntactic (unambiguous OBJ identification) and discourse-related (specificity, definiteness, relative animacy, prominence), as well as socio-dialectal, as Tippetts (2011) amply demonstrates.

Focusing on discourse now, discourse-related processing instructions (e.g., [make argument X prominent]) are part of the preverbal message, which instigates the lexical selection process (cf. Levelt 1989: 98-99; cf. § 2.1, ch. 1, this volume). These instructions about the assignment of prominence, itself an option exercised by the speaker, will then be carried out by the lexical choice. The chosen lemma in turn will include its morphological instructions – in the case of DOM whether the *a* marker is required or not.⁹ In other words, the lemma receiving the prominence

9 This reflects our theoretical position, which should be subjected to research into the processing of prominence – as yet an under-represented area in speech processing research.

instruction checks its grammatical function (i.e., OBJ) and, if SUBJ is already syntactically prominent (i.e., default TOP), then OBJ assumes its prominence (i.e., TOP₂) by other (e.g., prosodic and morphological) means, which in Spanish is the so-called personal *a*.

Indeed the morphology results indicate clearly that it is not enough for learners to be at the interphrasal stage – or even the interclausal S-BAR stage (stage 7 in Johnston's (1995) stages) – to be able to mark OBJ differentially. Both ROS and BEN, for instance, are at the top of the morphological schedule, and yet, having produced clear contexts for DOM, they are not using it. On the other hand JAN, like BEN, is able to mark S-BAR procedure as required for many subjunctive clauses in Spanish, but, unlike her, she does mark DOM. More surprisingly TAN, who is not at the S-BAR stage for morphology, is nevertheless able to mark DOM when the occasion arises. Her single occasion is valid evidence because she is also capable of topicalising OBJ.

So, in light of the similarity with Farley & McCollam's (2004) results considered in § 2, we may ask whether DOM is a necessary resource for processing subjunctive clause marking. The answer appears to be negative, because DOM is to some extent optional. Furthermore, DOM is just one of the structures requiring the S-procedure, and PT predicts that S-BAR procedure structures will appear after the emergence of *any* S-procedure structures (all learners in this study have acquired one or more S-procedure structures). Hence there is no contradiction within PT if a learner does not exhibit DOM and achieves subjunctive marking at the same time. Our results appear to support the specific implicational progression proposed in (19) within an overall developmental framework. The last step hypothesised in this progression awaits validation from a richer database.

On the overall evidence, it may be said that the S-procedure is a necessary resource for DOM to be marked, but not a sufficient condition, since discourse factors are also involved. This means that there must be a sufficient degree of control, or automatism, of more basic (e.g., lexical, phonological, and morphological) components for the learner to be able to integrate specific discourse-pragmatic information (cf. the discussion on the interface between morphological and syntactic development in § 4.3, ch. 1, this volume).

5. Conclusion

In this chapter we have engaged in an exploration of the acquisition of DOM in L2 Spanish and shown that current PT is capable of accounting for the optionality of its use in processing terms – that is, discourse-pragmatic information needs to be integrated online with semantic and syntactic information, which in turn requires morphological resources.

In terms of our research question – whether learners of L2 Spanish acquire DOM at the hypothesised developmental points in the PT schedules proposed in § 3 – the answer is positive. We have thus placed DOM within a developmental framework and indicated a specific implicational trajectory. Despite an important limitation in terms of data robustness, it is also clear that our results are compatible with those found by Guijarro-Fuentes (2011, 2012), in the sense of pointing to animacy as the default feature for marking DOM, as supported also by Tippetts' (2011) corpus work. They are also similar to Farley & McCollam's (2004), but our approach offers a detailed and multifaceted explanation for the phenomenon. With more robust, possibly longitudinal, data sets other acquisitional issues may be addressed, such as those surrounding DAT marking and clitics, as pointed out repeatedly in Montrul's work, the sociolinguistics of DOM in expatriate communities and its development in L1, as initiated by Rodríguez-Mondoñedo. Fully developmental studies are then desirable, and may clarify complex issues better than highly focused work on single phenomena.

Acquiring constituent questions in Italian as a second language

Camilla Bettoni and Giorgia Ginelli

University of Verona

1. Constituent questions

In Levelt's Model, the speaker must decide early whether the message to be conveyed is declarative, imperative, or interrogative (Levelt 1989: § 3.5), so indicators of mood are already present when the message generated in the conceptualizer reaches the formulator, and activates functional processing and positional processing (cf. § 2.1, ch. 1, this volume). In constituent questions the information the speaker wants to receive from the addressee is expressed by one (or more) question phrase (QP from now on) or word (QW), which fills the gap left by the focal constituent (or constituents) of the corresponding declarative sentence (Van Valin 2001: 185). We can tell that the interrogative phrase fills this gap by the fact that we cannot put another phrase of the same type in the same sentence. For example, the sentence in (1) is ungrammatical, because V can govern only one GF per type, with the exception of the ADJ function (Darlymple 2001: 11). Here the V *kiss* governs the OBJ function, and the presence of two OBJ arguments in the same sentence (*whom* and *the girl*) is ungrammatical. On the other hand, (2) is grammatical because each QP (*what* and *where*) fills the two required GFs (OBJ and OBL respectively) governed by the V *put*.

- (1) [John kissed the girl] *whom did John kiss the girl?
 (2) [John put the book onto the table] what did John put where?

Constituent questions are used to request some piece of new information, the focus of the sentence. Focality is not a prerogative of interrogatives, but in interrogatives FOC is obligatory (Choi 1999), and undergoes specific constraints. These vary cross-linguistically: all languages can satisfy the communicative need of requesting information, but different languages do it using different linguis-

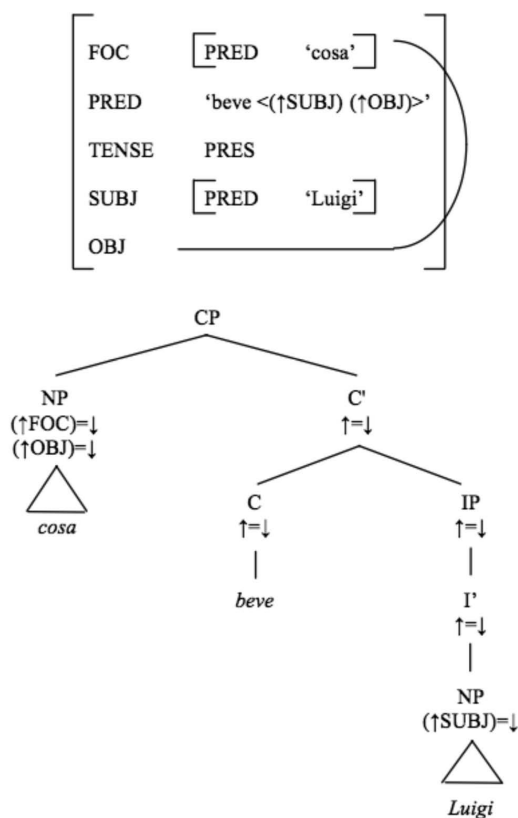
tic strategies. A derivational approach describes these strategies in terms of syntactic movement of QPs (cf. Chomsky 1977, as well as Cheng & Corver 2006; Rizzi 1997), and assumes a typological division between languages that formulate constituent questions by placing interrogative phrases in initial position (simple- or multiple-fronting languages), and those that leave them in the same position of their non-interrogative equivalents (in-situ languages). In the non-derivational framework of LFG, constituent questions are described as a particular kind of ‘filler-gap structures’ (Falk 2001; Kroeger 2004). These structures imply the existence of a missing element at the ‘gap’ position, which is the actual specific piece of information required, and of a ‘filler’ bearing two functions: (i) the DF FOC, associated with the interrogative phrase; and (ii) the GF, associated with the gap. In (2) above, for example, we can see the double status of the filler (*what*), bearing both the DF FOC, associated with the interrogative phrase, and the GF, in this case OBJ, associated with the gap. In LFG’s f-structure, abstract GFs and features try to capture universal syntactic principles that vary cross-linguistically at other levels of representation. Hence, the f-structure of constituent questions is independent from language-specific connotations like, for instance, word order: “since the pragmatic function of constituent questions is much the same in all languages, the functional structure of constituent questions in various languages is likely to be quite similar even when phrase structure configuration is very different” (Kroeger 2004: 171; cf. also Bresnan 2001: 45).

According to Mycock (2007), both these approaches – the derivational approach and the functional one – as previously formulated, fail to explain crucial aspects of constituent questions such as prosody. That is, they can capture only those principles that underlie the formation of constituent questions in languages like English that realize the focusing of QPs syntactically. However, in languages like Japanese, the QP appears in situ, that is, in the same position it occupies in the equivalent declarative sentence, and the focusing of the interrogative phrase is realized only prosodically. Because the DF FOC is not indicated at f-structure, or indeed at c-structure, Mycock (2007) maintains that constituent questions can be reduced neither to their word order nor to their f-structure, but must be described in terms of both the informational distribution at i-structure level, and the ways in which their i-structural status (i.e., FOC) is marked in syntax and/or in prosody. She demonstrates convincingly that the only universal feature of constituent questions is the focusing of the QP, and that what varies cross-linguistically is how it is focused: syntactically, and/or prosodically. For the theoretical development of incorporating within LFG the possibility of at least three other structures, besides a-structure, c-structure and f-structure, see Choi (2001), Falk (2001: 22ff), and more recently Dalrymple & Nikolaeva (2011).

2. Constituent questions in Italian

Most grammatical descriptions of Italian interrogatives indicate syntax as the key feature in the formation of constituent questions (e.g., Lepschy & Lepschy 1981; Fava 1995; Serianni 1996; Dardano & Trifone 1997; Salvi & Vanelli 2004; Schwarze 2009). This does not mean that prosody does not play a role in Italian, but only that the most important aspect of Italian constituent interrogatives is syntax. Leaving prosody for future investigations, we will deal here with the syntactic strategies that in Italian allow for the focusing of the new information enclosed in QPs.¹

(3) *F-structure and c-structure of cosa beve Luigi? (what does Luigi drink?)*



1 In his taxonomy of Focus, Dik's (1997: 331ff.) distinction between Questioning Focus and other types – notably Completive (New) Focus – is insightful.

Typologically, Italian is a headmarking language located towards the less configurational end of the typological spectrum (cf. fig. (1), part II, this volume). Although its canonical order is SVO, word order is relatively free, regulated by discourse and pragmatic choices (cf. § 3.1, ch. 3). In constituent questions, word order is marked, because the QP appears initially in the sentence rather than in the position typical of its GF. The initial position then, in interrogatives, is associated with the DF FOC. Because Italian is a language with syntactic focusing, FOC is grammaticalized and indicated at f-structure, where it bears also the GF specified by the argument list of the verbal predicate. Consider the sentence in (3), for example, with the illustration of its f- and c-structures. Here word order is marked in two ways. First, the interrogative phrase bearing the FOC function – co-referential with the OBJ gap after the V *beve* – appears clause-initially, that is, not in the position associated with its GF. Secondly, SUBJ appears postverbally, that is, not in its canonical preverbal position.

The only exception to the markedness of word order in interrogatives is when the focal QP itself bears the GF SUBJ, as in (4). In this case, SUBJ is questioned in situ, clearly showing its special status as the only core GF which, in LFG, is also a DF – FOC, in this case. Also in English SUBJ is the only constituent that is questioned in situ and does not require AUX, as in *who drunk the wine?* vs *what did he drink?* (cf. Falk 2006 for a different approach to the notion of SUBJ in LFG which reassesses this apparent anomaly and discusses thematic roles, a-structure, GFs, and the mappings between them).

- (4) chi beve il vino?
 who.SUBJ drink-3.SG the wine?
 [who drinks the wine?]

Word order of Italian constituent questions can be marked in yet a third way if the interrogative sentence is itself pragmatically marked, and there is a topicalised constituent in first position. Consider the sentence in (5), for example, where also TOP=OBJ appears in a noncanonical position, besides FOC=OBL_{GOAL} and SUBJ. The complexity of this structure is further increased by the fact that in Italian the topicalization of OBJ requires the use of a clitic as morphological marker of OBJ onto V, as shown in §§ 3.1-3.2, chapter 3, this volume.

- (5) il vino a chi lo offre Luigi ?
 the wine-TOP.SG.MASC to whom-FOC_{OBL} it-CL.ACC.SG.MASC offers Luigi-SUBJ
 [the wine to whom does Luigi offer it?]

3. A developmental hypothesis

For the development of Italian constituent questions by L2 learners, we take as our starting point PT's hierarchy presented as the Prominence Hypothesis in (34)-(35), § 4.2.1, chapter 1, this volume, and implemented for Italian declarative sentences in (23), § 3.1 in chapter 3. Our hypothesis is illustrated by the hierarchy in (6).

(6) *Developmental stages hypothesised for Italian L2 syntax based on the Prominence Hypothesis: interrogatives (after Bettoni & Di Biase 2011)*

STAGE	STRUCTURE	EXAMPLE
NONCANONICAL WORD ORDER	TOP _i FOC _{WH} - Cl _i -AUX V- <i>to</i> _i SUBJ?	<i>i libri a chi li ha dati Pia?</i> [the books, to whom has Pia given them?]
	TOP _i FOC _{WH} - Cl _i -V SUBJ?	<i>i libri a chi li dà Pia?</i> [the books, to whom does Pia give them?]
	FOC _{WH} - VS	<i>che cosa mangia Piero ?</i> [what eats Piero]
XP _{DF} CANONICAL WORD ORDER	FOC _{WH} - SV	<i>che cosa Piera mangia ?</i> [what Piera eats ?]
		<i>che cosa mangia ?</i> [what (he) eats ?]
CANONICAL WORD ORDER	SV FOC _{WH} -	<i>Paola mangia cosa? [Paola eats what?]</i>
		<i>Paolo va dove? [Paolo goes where?]</i>
LEMMA ACCESS	single words	<i>cosa? quando? [what? when?]</i>
	formulas	<i>come stai? [how are you?]</i>

Studies within the PT framework on the acquisition of Italian simple, pragmatically neutral, minimally presuppositional declarative sentences (e.g., Di Biase & Kawaguchi 2002; Di Biase & Bettoni 2007; Bettoni, Di Biase & Ferraris 2008; Bettoni, Di Biase & Nuzzo 2009; cf. also § 3.1, ch. 3, this volume) report that initially learners can produce only sentences with canonical word order whose structure has a fixed correspondence between arguments (such as agent, patient) and c-structure. Also questions go through this canonical word order stage. However, soon after learners have the lexical means to convey an interrogative message and the morphological means to distinguish between Ns and Vs, the canonical word order stage is quickly overtaken. In fact, constituent questions, as we have seen, are sentences marked pragmatically, whose very essence is the focalisation of QP.

Thus, at this early stage, fronted QP FOC is followed by canonical word order. Having yet to learn to assign GFs to verbal arguments independently of position, learners do not master the means to free up the canonical order sequence, as required by Italian constituent questions. The outcome is then ungrammatical, as in (7)-(8), when SUBJ (expressed referentially or pronominally) is in preverbal position without being questioned (that is, when FOC \neq SUBJ).

- (7) *che sport Alberto fa?
 which sport Alberto do-3.SG?
 [which sport does Alberto do?]
- (8) *dove tu trovi questi fiori?
 where you find-2.SG these flowers?
 [where do you find these flowers?]

On the other hand, the outcome is grammatical in either of two cases: when the sentence is null SUBJ, as in (9), or the QP in initial position bears itself the SUBJ function, as in (10) or (4) above.

- (9) cosa studia?
 what study-3.SG?
 [what does (she) study?]
- (10) chi ha questa borsa?
 who have-3.SG this bag?
 [who has this bag?]

At this same XP canonical order stage, as well as fronted focal QWs or QPs, learners can also produce topical words or phrases. The resulting sentences can then vary in terms of grammaticality, depending on the GF associated with TOP. When it bears the ADJ function, which is an ungovernable noncore function, the outcome is ungrammatical, as in (11)-(12), or grammatical, as in (13)-(14), not because of the presence of ADJ (which in the target language does not require any adjustment in the following sequence), but for the same reasons we have just seen regarding the examples in (11)-(14).

- (11) *adesso che sport Alberto fa?
 now which sport Alberto do-3.SG?
 [now which sport does Alberto do?]
- (12) *in giardino dove tu trovi questi fiori?
 in the garden where you find-2.SG these flowers?
 [in the garden where do you find these flowers?]

- (13) all'università cosa studia?
 at the university what study-3.SG?
 [at the university what does (she) study?]
- (14) questa mattina chi ha comprato il pane?
 this morning who have-3.SG bought the bread?
 [this morning who has bought the bread?]

On the other hand, if TOP bears the OBJ function and its coreferential OBJ clitic marker on V is missing, the TOP function is left unmarked and the outcome is ungrammatical, as in (15) – cf. Bresnan's (2001: § 4.8) notion of 'functional uncertainty'.

- (15) *la cartolina dove hai comprato?
 the postcard where have-2.SG bought?
 [where have you bought **the postcard**?]

In this example, even if the sentence is ungrammatical, the meaning is quite clear lexically, just as it is clear with topicalised ADJ. This is because the listener is unlikely to take the inanimate OBJ *postcard* as SUBJ doing the action of buying. Likewise, when TOP is ADJ, the fact that this GF is usually expressed by a circumstantial adverb or a PP rules out the possibility for the listener to take it as SUBJ, which by default occupies the first position in canonical order. Problems of comprehension, however, do arise when both OBJ and SUBJ have animate referents. In this case, if the coreferential clitic marking the first constituent as OBJ is missing on V, the reader will indeed take it as SUBJ (cf. (19)–(21), ch. 3, this volume).

One stage further up in the developmental path, at the noncanonical word order stage, having learned the necessary morphology to assign GFs to constituents (cf. § 4.2.1, ch. 1, this volume), learners can now mark them by means other than position, and thus free up canonical word order. By placing SUBJ in postverbal position, as in (16), they will codify grammatical constituent questions also when there is a pronominal or referential SUBJ and the fronted QP bears GFs other than SUBJ.

- (16) che cosa compra Paolo?
 what buy-3.SG Paolo?
 [what does Paolo buy?]

Also at this stage, like at the previous stage, besides focalising QP, learners may wish to topicalise constituents with a variety of GFs. When TOP is ADJ, as in (17), the

sentence is grammatical because no further adjustment is required. On the other hand, when TOP is OBJ, should the missing coreferential OBJ clitic leave the GF unspecified, the sentence would be ungrammatical, as in (18). The agreement between TOP and the clitic obviously increases the cognitive load of the production of a topicalised constituent question. Thus, although technically still within the noncanonical word order stage, our hypothesis is that the sentence in (19) will appear after that in (18).

- (17) ieri cosa ha fatto Alberto?
 yesterday what have-3.SG done Alberto?
 [yesterday what did Alberto do?]

- (18) *i libri dove compra Pino?
 the books where buy-3.SG Pino?
 [where does Pino buy **the books?**]

- (19) i libri dove li compra Pino?
 the books-MASC.PL where them-MASC.PL buy-3.SG Pino?
 [where does Pino buy the books?]

In this type of sentence, not only FOC QP, but also both the TOP_{OBJ} and SUBJ core functions appear in noncanonical position. The progress now occurs not only because learners manage, first, to identify the DF and GF of each NP, but also because they manage, secondly, to mark OBJ morphologically onto V with the clitic coreferential with TOP – cf. (5) above. At this same stage there is one further step to go in the learning process. When V is inflected analytically, Italian requires that the past participle agree in number and gender with the clitic. This adds a further burden to the online production of the sentence. Our hypothesis, then, is that there will be yet another step within the stage when learners first produce ungrammatical sentences with a default masculine singular past participle, as in (20), and finally fully grammatical topicalised interrogative sentences, as in (21).

- (20) i libri dove li ha *comprato Pino?
 the books-MASC.PL where them-MASC.PL have-3.SG bought-MASC.SG Pino?
 [where has Pino bought the books?]

- (21) i libri dove li ha comprati Pino?
 the books-MASC.PL where them-MASC.PL have-3.SG bought-MASC.PL Pino?
 [where has Pino bought the books?]

4. The study

In order to verify the developmental hierarchy hypothesized in § 3, we analyse cross-sectional data pertaining to 12 learners with different levels of competence in Italian. All learners are European students in their early twenties attending Italian L2 courses at the University of Verona: five of them (Ve, Pe, Jh El, Me) are Czech, two (Cr, La) are German, and one each among Lu, Ni, Ju, Ma, and Ev is respectively Dutch, English, Russian, Spanish and French – their names of course having all been changed. The study also includes one native speaker control (Ga). Because most of the structures tested are optional, this inclusion allows for comparison between the learners' and the native speaker's production in the same situations. All subjects were recorded in March and April 2008.

The data elicitation tasks for this study are partly inspired by those used for English questions (cf., e.g., Pienemann 1998: 280; Keatinge 2008), and partly specifically devised for Italian, which, unlike English, is a null-SUBJ language. Learners are prompted to produce interrogative sentences in order to gather information on two different items or events (e.g., two mysterious objects, two types of weddings). Because SUBJ position is of great relevance in the grammatical codification of Italian constituent questions (cf. §§ 2-3), the presence of two items encourages the use of explicit SUBJ rather than the more common null SUBJ. In order to make sure that learners do use it, they are not allowed to collect all the information they require to play the game first on one item and then on the other, but must ask questions about them alternatively, thus specifying an item each time. In our study, five tasks were specifically targeted to elicit constituent questions, other tasks served as distractors.

Out of the whole corpus thus collected, we analyse here constituent questions encoded in full sentences with lexical Vs whose a-structure maps canonically onto f-structure, such as *dare* (give) and *comprare* (buy). In other words, we do not consider passives or sentences with unaccusative and so-called exceptional Vs (Pinker 1984), as well as copular and presentative sentences (for these 'nonverbal predicates', cf. Kroeger 2005: ch. 10). Furthermore, in this study we consider only sentences in which the focal element bears an argument GF or the nonargument GF ADJ. This means that we leave for future analysis those with *perché* ('why') and *come* ('how'), which may involve subordination (regarding *why* in English L2, cf. § 3.1, ch. 2, this volume). Finally, formulas such as *quanto costa?* ('how much is it?') have been discounted. Altogether, we analyse 372 constituent questions, corresponding to an average of circa 29 for each of the 12 learners, compared to 33 by Ga, the native speaker of Italian.

5. The analysis

The empirical support for the interlanguage development of Italian constituent questions hypothesised in (6) is presented cross-sectionally in (22).

(22) *Cross-sectional study of the development of Italian constituent questions based on the Prominence Hypothesis: interrogatives*

STRUCTURE	VE	PE	LU	LA	JU	CR	JH	NI	EL	MA	EV	ME	GA
TOP _i QP Cl _i -AUX V- <i>to_i</i>										-1	-2	1	1
TOP _i QP Cl _i -V											2		
TOP QP V SUBJ							1	3					2
QP V SUBJ	7	23	6	6	6	12	14	15	8	6	1	13	8
TOP QP V	1								1				3
QP SUBJ V	12	6	4	4	5	4	6						
QP V	7	7	4	16	12	16	9	12	21	20	11	32	16
(SUBJ) V QP	1			1					1			2	3

As shown in (22), in-situ questions belonging to the canonical word order stage are few and produced by only some of the learners, as in (23)-(24), as well as by Ga, the native speaker.

(23) Ve tu preferisci quale musica?
 you prefer-2.SG which music?
 [which music do you prefer?]

(24) La resta all'università a che ora?
 stays-3.SG at university at what time?
 [until what time does he stay at university?]

All our learners have also reached the XP canonical word order stage in so far as they produce 8 or more structures where QPs are clause-initial and fronted, followed by canonical word order. There is also ample evidence that all learners can use a wide range of QWs and QPs – even Pe, for example, as shown in (25). As a matter of fact, in this regard, he displays a range of forms that is already similar to Ga's, the native speaker.

- (25) Pe
- a. dove tu trovi questi fiori?
where you find-2.SG these flowers?
[where do you find these flowers?]
 - b. che cosa fa in casa a sera?
what do-3.SG at home in the evening?
[what does (he) do at home in the evening?]
 - c. chi ha questa borsa?
who has-3.SG this bag?
[who has this bag?]
 - d. quanto tempo tu conosci questi amiche amici?
how long you know-2.SG these friends?
[how long have you known these friends?]
 - e. qualche [quale] artista preferisce?
which artist prefer-3.SG?
[which artist does (he) prefer?]

There is, however, significant variation among the 12 learners in the distribution of the different structures belonging to this XP canonical word order stage, concerning two factors: (i) the use of SUBJ vs null SUBJ; and (ii) the addition of a topicalised element. First, the less advanced learners – in particular Pe, Lu and Ve – use null SUBJ quite sparingly, and overuse instead pronominal SUBJ in pragmatically unmotivated contexts, as exemplified in (25)–(26): Ve 13 times out of 20, Pe and Jh 6 times out of 13 and 15 respectively, Lu 4 out of 8. In (25b), Ve even uses both a pronominal and a referential SUBJ; needless to say, in this latter case preverbal SUBJ is ungrammatical.

- (25) a. Ve quanto tempo tu passi con Marco?
how much time you spend-2.SG with Marco?
[how much time do you spend with Marco?]
- b. quale ora lei Daniela va dormire?
at what time she Daniela go-3.SG sleep?
[at what time does Daniela go to sleep?]
- (26) Pe quanto tempo fa loro sono incontrato?
how long ago they be-3.PL met?
[how long ago have they met?]

As learners progress, they display a more frequent use of null SUBJ. So, Cr and La already show some progress in two ways: their figures for null SUBJ are high-

er (they both use null SUBJ 4 times out of 16), and their use of SUBJ, although still ungrammatically preverbal, is more pragmatically justified, and referential rather than pronominal, as in (27). Finally, and most proficiently, Ni, El, Ma, Ev, and Me use null SUBJ exclusively, like Ga does. This results in grammatical sentences, as in (28)

- (27) Cr quando Daniela deve iniziare lavorare?
 when Daniela must-3.SG start work?
 [when must Daniela start to work?]

- (28) El quante domande devo fare?
 how many questions must-1.SG make?
 [how many questions must (I) make?]

Still within this stage, the second significant variation among the learners concerns the addition of a topicalised element. Once learners are able to front FOC, the addition of a topicalised constituent in the prominent clause-initial position would not seem very costly, provided TOP bears the ADJ function. Yet only 2 of our 12 learners topicalise ADJ, and do so once each, compared to Ga who does it 3 times. The grammaticality (or ungrammaticality) of the full sentence with this addition is independent of fronted ADJ, and depends exclusively on what follows it: in our corpus the sentences are null SUBJ, as in (29)-(30), so the outcome is targetlike, even for Ve, who is the weakest learner.

- (29) Ve dopo lavorare che cosa fe [fa]?
 after work what do-3.SG?
 [after work what does (she) do?]

- (30) El adesso quale hai in mano?
 now which have-2.SG in hand?
 [now which one do you have in the hand?]

Although the use of a topicalised ADJ is rare among the learners, it is not altogether rare for some of them to preface their sentences with an external element. We do not consider this element a DF TOP, however, because it is independent of the structure that follows. In fact it is then repeated within the sentence, where it bears a variety of GFs, such as SUBJ in (31)-(34), OBJ in (32), and ADJ in (33). Notice that the following sentence in (31) repeats SUBJ, whereas in (34) null SUBJ follows.

- (31) Lu Roberto a che ora Roberto mangia in la sera?
 Roberto at what time Roberto eat-3.SG in the evening?
 [Roberto, at what time does Roberto eat in the evening?]

- (32) Ni i fiori chi ha dato i fiori a te?
 the flowers who have-3.SG given the flowers to you?
 [the flowers, who has given the flowers to you?]
- (33) Me sul lago cosa hanno fatto sul lago?
 on the lake what have-3.PL done on the lake?
 [on the lake, what have they done on the lake?]
- (34) Ve Roberto cosa fe [fa]?
 Roberto what do-3.SG?
 [Roberto, what does he do?]

Further up the developmental hierarchy, at the noncanonical word order stage, learners acquire functional assignment, and can thus free up canonical word order and place SUBJ postverbally. As (22) shows, all our learners have also reached this stage – even Ve and Pe, who still have problems with the lexicon and verbal inflection, as exemplified in (35), where the latter first uses *qualche* instead of the targeted interrogative adjective *quale*, and then variously attempts *preferi*, *preferà* and *preferiscia* instead of the targeted third person singular *preferisce* ('(he) prefers') of the V *preferire*.

- (35) Pe a. qualche cibo preferi preferà Marco?
 what food prefer-3.SG Marco?
 [which food does Marco prefer?]
- b. qualche musica preferiscia Marco?
 what music prefer-3.SG Marco?
 [which music does Marco prefer?]

However, although all learners can place SUBJ postverbally with what appears to be nativelike online processing, if we look at the distribution of SV and VS, we can see that some learners are more advanced than others. That is, as we move from the left towards Ga on the right in (22), we can see that more beginner learners alternate postverbal with preverbal SUBJ; more advanced learners do it less, in favour of postverbal SUBJ, thus becoming more accurate; until they no longer do so, with Ni, El, Ma, Ev, and Me producing only grammatical sentences.

Still within the noncanonical word order stage, like at the previous stage, some questions are preceded by topicalized constituents. These can bear the ADJ function as in (36), and do not imply further linguistic constraints in the grammatical codification of the sentence; the outcome is thus grammatical.

- (36) Ni per pranzo cosa ha fatto Maurizio ieri?
 for lunch what have-3.SG done Maurizio yesterday?
 [for lunch what has Maurizio done yesterday?]

On the other hand, several constraints are operative when the topicalized constituent bears the OBJ function. As we have seen in §§ 2-3, in order to codify the sentence correctly learners must be able, first, to identify the GFs of all NPs and, secondly, to unify the number and gender features of TOP with the clitic, and the past participle if V is in analytical form. Among our 12 learners, three have reached this highest level within the last stage in the developmental hierarchy. Their figures are not robust, but then also Ga produces only one OBJ topicalisation among his 30 questions. Coupled with the fact that this is a cross-sectional study with only 12 learners, this means that our data does not allow us to have a clear developmental picture of the full range of variation within this last stage. First, all 6 occurrences of OBJ topicalisation are null-SUBJ sentences. This is quite appropriate, but it can tell us nothing about the placing of SUBJ. All we can say in this regard is that, if we look at the figures of VS and SV sequences for these three learners in all their production, we notice that they always place SUB accurately in postverbal position. Secondly, all three learners always mark the TOP function with formally accurate clitics, although Ma seems somewhat hesitant when she produces *i* and *le* before getting *li* right, as in (37). This means that we do not have enough variation in the data to prove that the default masculine singular *lo* clitic might initially be overused.

- (37) Ma i fiori chi i le ti li ha *dato?
 flowers-MASC.PL who to you them-MASC.PL have-3.SG given-MASC.SG?
 [the flowers who has given them to you?]

Thirdly, Ma and Me produce only one OBJ topicalisation each, both of them with V in analytical form. Having no synthetic form to compare them with, in order to prove our hypothesis that the number and gender agreement will be marked last on the past participle, it is irrelevant to know that Me marks the past participle correctly in (38), and Ma does not, using *dato* instead of *dati* in (37).

- (38) Me Elisa dove l' hai conosciuta?
 Elisa-FEM.SG where her-FEM.SG have-2.SG met-FEM.SG?
 [Elisa where have (you) met her?]

The only learner providing evidence in support of this hypothesis of a step within a stage is Ev, who produces four OBJ topicalisations: two with V in synthetic form

which are fully accurate, as in (39a), and two with V in analytical form which are accurate except for the missing agreement on the past participle, as in (39b).

- (39) a. Ev Roberta da quanto tempo la conosci?
 Roberta-FEM.SG since how long her-FEM.SG know-2.SG?
 [Roberta for how long have you known her?]
- b. i fiori dove li ha *messo nella casa?
 flowers-MASC.PL where them-MASC.PL have-3.SG put-MASC.SG in the house?
 [the flowers where has (he) put them in the house?]

6. Conclusion

In this exploratory study of the development of constituent questions in the Italian interlanguage of L2 learners, our empirical cross-sectional data fully supports the hypothesis presented in § 3. Despite the great care taken in devising suitable elicitation tasks, numbers of occurrences for some of the structures, especially at the highest stage, are low, but this is not surprising, given both their infrequent occurrences also in native speakers' production, and their wide range. This range involves, on the one hand, the presence or absence of SUBJ, and its position when present; and on the other, several different GFs associated with focusing question constituents and topicalisations.

To summarise our findings, we have seen that all our learners are lexically ready to produce constituent questions in so far as even the weakest among them produce as wide a range of QWs and QPs as the native speaker does. Syntactically, all our learners can focus the question constituent by fronting it. The prime reason for this is that fronted FOC is the Italian way of marking interrogatives, and hence the most pervasive QP position in the input. This also explains why only a few learners in our data appear to traverse the canonical word order stage by producing in-situ questions. In this regard, we draw the reader's attention to a similarity with German L2 in chapter 9 regarding the quick way in which learners appear to disentangle canonical word order thanks to the strength of the FOC position in constituent questions. This may very well show that canonical word order is not as relevant in questions as in declaratives. Or indeed it may even turn out that declaratives and constituent questions have different canonical word orders.

Besides placing QPs in initial position, all our learners can also place SUBJ postverbally. This means that all of them have reached the noncanonical stage, and are therefore – stagewise – quite advanced. However, if we look at accuracy beyond emergence, we notice fair differences among the learners in terms of the proportion of structures which are pragmatically justified and syntactically accurate, and

those which are not. As we move along the developmental continuum, on the one hand, pragmatically, learners gradually curb their tendency to overuse redundant pronominal SUBJ until their null-SUBJ structures become as frequent as those produced by the native speaker. On the other, syntactically, when referential SUBJs are motivated by discourse or pragmatic reasons, learners gradually reduce their preverbal ungrammatical SUBJs until they produce only targetlike postverbal grammatical ones.

The addition of a topical element in front of the focal QP is an altogether rare occurrence in both the learners' and the native speaker's data. When learners do use it, no progress is discernible among them as long as the added topical element is associated with the ADJ GF and requires no further formal constraints. On the other hand, when the topic is associated with the OBJ function and costly constraints are required, we have a clear indication that, among our 12 learners, only three of them have moved ahead of the others and produce at least one OBJ topicalisation, although with some variation in the accuracy of morphological agreements.

We can conclude that the hypotheses we put forward in (22) are fully supported by the empirical data of this small cross-sectional study. They would have been falsified if they had contradicted the implicational relationship among the predicted stages. But predictions have turned out positively, thus strengthening the case for PT's universal hierarchy presented in (35), chapter 1, this volume. As well as generally confirming the hierarchy for the development of constituent questions in a nonconfigurational language like Italian, this study shows specifically that in such a language it may be easier to move from one stage to the next in terms of emergence of syntactic structures, than to progress in terms of accuracy by either pragmatically discerning when to use alternatives (i.e., null SUBJ vs full SUBJ; fronted vs in-situ QPs) or morphologically increasing the range of inflections within a stage (i.e., default vs nondefault number and gender markers).

Acquiring V2 in declarative sentences and constituent questions in German as a second language

Louise Jansen* and Bruno Di Biase**

*Australian National University, **University of Western Sydney

1. Introduction

German is a well-studied language within the PT framework and its predecessor, the Multidimensional Model (cf. Meisel, Clahsen & Pienemann 1981; Clahsen, Meisel & Pienemann 1983; Pienemann 1981, 1984, 1989, 1998; Clahsen 1984; Ellis 1989; Boss 1996, 2004; Jansen 2008; Baten 2013). However, no study has yet focused on constituent questions specifically for German, nor tested the Topic Hypothesis (Pienemann, Di Biase & Kawaguchi 2005) for this language or indeed its updated version, the Prominence Hypothesis, proposed in chapter 1, this volume. Furthermore, no study has yet conceptually and empirically compared the development of interrogative and declarative sentences, as proposed in this volume, in order to assess whether this generates any significant theoretical consequences. This chapter aims at filling these gaps on the basis of data collected for a study of the development of German syntax by Jansen (2008). Since the 2008 corpus includes content questions, we have the opportunity to investigate their development and compare it with that of declaratives in those learners who do produce both. In what follows, we will first present some relevant principles of German syntax, and hypothesize the developmental hierarchies for declaratives and interrogatives according to the Prominence Hypothesis (§ 2). There follows a brief consideration of the literature on the acquisition of German word order (§ 3) reporting, in particular, on results of two historically significant studies conducted within the PT traditional background (i.e., ZISA/Clahsen, Meisel & Pienemann 1983; Pienemann 1989). Then, in (§§ 4-5), the methodology of this study and the results of the current investigation on content questions will be presented, along with those for declaratives in Jansen's (2008) study, as guided by two main research questions:

- Do learners of German L2 follow the developmental schedule according to the Prominence Hypothesis?
- Does noncanonical word order develop in both constituent questions and declaratives at the same time and in the same way?

Finally, in the discussion (§ 6), we will summarise and interpret the results of all three studies (i.e., ZISA, Pienemann, and our own), and conclude that learners of German L2 (a) indeed follow the PT schedule, (b) clearly begin their production of noncanonical word order in constituent questions much earlier than in declaratives, and (c) appear to produce questions with noncanonical word order categorically at their onset, whereas with declaratives the path from emergence to acquisition seems much more gradual.

2. German syntax and the Prominence Hypothesis

There are two main peculiarities in German word order: German is a so-called ‘verb-second’ (V2) language, and its canonical word order differs according to whether the lexical V is inflected or not. Let us look at these in turn.

First, in the main clauses of a V2 language the inflected V obligatorily occupies the second position in c-structure. A corollary of V2 is that XP SVX is ungrammatical in German, as in (1), because this would force the V into third position.

- (1) *heute ich spiele Tennis
 XP SUBJ V OBJ
 today I play tennis

Hence, if any constituent bearing a function other than SUBJ (whether ADJ or argument function) occupies the sentence-initial position usually associated with SUBJ/TOP, SUBJ must follow the finite V, as in (2a-b).

- (2) a. heute spiele ich Tennis
 TOP_{ADJ} V SUBJ OBJ
 today play I tennis
 [today I play tennis]
- b. Tennis spiele ich heute
 TOP_{OBJ} V SUBJ ADJ
 tennis play I today
 [tennis I play today]

Second, as pointed out by Zobl (1986), canonical word order in German is SVO in main clauses with lexical Vs only, whereas the lexical V is clause-final in main clauses with AUX-V or MOD-V, as shown in (3a-b).

- (3) a. ich spiele Tennis
 SUBJ V OBJ
 [I play tennis]
- b. ich habe Tennis gespielt
 SUBJ AUX OBJ V
 [I have played tennis]

Thus, learners with a V-last L1, such as Turkish may first hypothesize that the V-last they hear in the input is canonical for German (Zobl 1986) as found, for instance, by Haberzettl (2005). We will come back to this point in (§ 3) below.

With regard to constituent question formation, German is a fronting language,¹ which means that FOC, the discourse function associated with question words or phrases (QP) is expressed, syntactically, in sentence-initial position (Mycock 2007). Parallel to the DF TOP (for a brief outline of these categories, cf. § 2.2, ch. 1, this volume) in declarative sentences, the DF FOC expressed in the QP can be linked to any GF, whether ADJ as in (4a), or argument as in (4b).

- (4) a. wann spielst du Tennis ?
 FOC_{ADJ} V SUBJ OBJ
 when play you tennis
 [when do you play tennis?]
- b. was spielst du heute ?
 FOC_{OBJ} V SUBJ ADJ
 what play you today
 [what do you play today?]

Thus the V2 rule applies to both topicalised declaratives and constituent questions. That is, SUBJ is postverbal whenever TOP or FOC, bearing any other GF, is in sentence-initial position.

Let us now hypothesise how learners acquire the German V2 rule in declaratives and constituent questions – or, in other words (i.e., those used in this volume), how they learn to move beyond the German SVO/SOV canonical word

1 No distinction is made here between single-fronting and multiple-fronting languages (cf. Mycock 2007).

order. Based on the universal progress outlined with the Prominence Hypothesis in chapter 1, this volume, we illustrate the development of German declaratives and constituent questions in (5).

(5) *Developmental stages hypothesised for German L2 syntax based on the Prominence Hypothesis: declaratives and constituent questions*

STAGE	DECLARATIVE SENTENCES		CONSTITUENT QUESTIONS	
	STRUCTURE	EXAMPLE	STRUCTURE	EXAMPLE
NONCANONICAL WORD ORDER	TOP V S	<i>Tennis spiele ich heute</i> [tennis play I today]	FOC V S	<i>was spielst du heute?</i> [what play you today?]
		<i>heute spiele ich Tennis</i> [today play I tennis]		<i>wann spielst du Tennis?</i> [when play you tennis?]
XP _{DF} CANONICAL WORD ORDER	TOP S V	<i>Tennis ich spiele heute</i> [tennis I play today]	FOC S V	<i>was du spielst heute?</i> [what you play today?]
		<i>heute ich spiele Tennis</i> [today I play tennis]		<i>wann du spielst Tennis?</i> [when you play tennis?]
CANONICAL WORD ORDER	S V O	<i>ich spiele Tennis</i> [I play tennis]	S V O ?	<i>du spielst was?</i> [you play what?]
	S O V	<i>ich Tennis spiele</i> [I tennis play]	S O V ?	<i>du was spielst?</i> [you what play?]
	S V	<i>ich spiele</i> [I play]	SV ?	<i>du spielst?</i> [you play?]
LEMMA ACCESS	formulas	<i>ich heiße Helmut</i> [my name's Helmut]	formulas	<i>wie geht's?</i> [how are you?]
	single words	<i>wir spielen</i> [we playing]	single words	<i>was spielen?</i> [what playing?]

For declaratives, we hypothesise that the German canonical word order, SVO or SOV, will emerge first. Topicalization followed by canonical order will come about next, followed finally by noncanonical word order. For constituent questions, the developmental hypothesis is parallel to that of declaratives. That is, the single question word stage is followed by a stage where an SVO or SOV order is used with in-situ QP FOC. The following stage will see a QP FOC in first position, followed by canonical word order. This is not standard German, but L2 learners are hypothesized to go through this stage regardless of whether their L1 exhibits V2 or not (cf. Pienemann & Håkansson 1999; Pienemann, Di Biase, Kawaguchi & Håkansson 2005). Finally noncanonicity will emerge.

3. Literature review

Before dealing with the development of questions we should briefly refer to the development of German canonical word order in L2 learners. As alluded to above, in her longitudinal study of four L2 German child learners, two of whom with Turkish L1, Haberzettl (2005: 159) finds that in their early data samples finite Vs are predominantly in final position; from this she concludes that her results for the two Turkish learners do not support PT. On the other hand, data of an adult Turkish learner from an earlier longitudinal study by Klein & Carroll (1992: 162–172), re-analysed by Schwarz & Sprouse (1994), shows that inflected V forms appear systematically in second position, and uninflected forms in final position. Furthermore, in their longitudinal analysis of an Italian learner of German L2, Klein & Carroll (1992: 131) observe “some degree of uncertainty”, with both SVO and SOV occurring in the same utterances. How do all these findings affect our proposed schedule in (5)? Contrary to Haberzettl’s interpretation, her data does not contradict PT because, unlike past interpretations of SVO as the universal canonical order, PT now hypothesises language-specific canonical orders (cf. § 4.2.1, ch. 1, this volume), and for German we hypothesise both the SOV and SVO orders. Neither, of course, do any of the Haberzettl’s, Klein & Carroll’s and Schwarz & Sprouse’s findings contradict PT’s Developmentally Moderated Transfer Hypothesis (Pienemann, Di Biase, Kawaguchi & Håkansson, 2005) because German canonical order includes both the L1 Turkish SOV and the Italian SVO.

With regard to the development of questions, to our knowledge there are no studies focusing specifically on their acquisition in L2 German, although two historically significant studies do include relevant data. These are (a) the ZISA cross-sectional study on the naturalistic acquisition of German by 45 learners from Romance language backgrounds (Meisel, Clahsen & Pienemann 1981; Clahsen, Meisel & Pienemann 1983), and (b) a longitudinal study on the instructed acquisition of German by three university learners from an Australian English background by Pienemann (1987, 1989, 1998). The explanatory frameworks applied to these two studies are different from the one proposed here. Whereas this volume uses LFG, which is a nonderivational approach, studies (a) and (b) assume a grammatical framework that is derivational (transformational), and hence group what for us is the noncanonical word order of declaratives and questions, emerging at the upper stage of the developmental path, under the category of ‘inversion’. Next we summarise the main results of these studies.

In (6) below we present cross-sectional data for the 25 learners, among the 45 investigated in the ZISA project, who produce constituent questions. Our table is constructed on the basis of two ZISA tables in order to facilitate the comparison between questions and declaratives, and further reorganised in order to make it more reader friendly in the context of this volume. In this connection, a couple of

points need to be raised. First, unlike all studies in this volume and Jansen's presented below, the ZISA analysis includes sentences with copula (cf. examples (2) and (7) in Clahsen, Meisel & Pienemann 1983: 128-129). Second, the ZISA study only includes the 'inversion' data, expressed as ratios of the application of the SUBJ-V inversion rule. So, '1' means that 'inversion' is categorically applied (100% of the times), 0.25 (or other fractions) means that the rule is applied 25% (or n%) of the times; and '0' that it is never applied (0%); an empty cell means no context for the rule application is produced. Because only obligatory contexts were included originally, we assume that nonapplication of 'inversion' (marked as 0 in the table) represents the production of structures of our XP canonical word order stage. Note also that the ZISA tables distinguish whether learners produce five or more contexts for 'inversion', or less than five. In the latter case the ratio is represented in brackets in (6) below as in the original ZISA tables. Finally, we maintain the original division in three columns and the original labels, but we add a row with

(6) *Application of the SUBJ-V inversion rule (i.e., production of structures at the noncanonical word order stage) in questions (first column) and declaratives (last three columns) – Reconstructed from Clahsen, Meisel & Pienemann (1983: 130-134, 145)*

ORIGINAL LABELS	WH-QUESTION	ADVERB	TOP1	TOP2
OUR LABELS	FOC _{WH}	TOP _{ADVERB}	TOP _{NP}	TOP _{SUB. CLAUSE}
LEARNERS				
JANNI	1	1	1	1
AGOSTINO	1	1	1	1
PINTOS	1	1	(1)	(1)
BENITO	1	1	1	0
EDUARDO	1	.94	(1)	(1)
MICHELE	1	.92	(1)	(1)
PIETRO	1	.88	(1)	.83
PEPE	1	.84	(1)	(1)
MARIA I.	1	.70	(1)	1
GIOVANNI	.67	1	1	(.50)
MARIA S.	(.33)	1	(1)	1
CETTINA	(1)	.40		
CARLO	1	.23	0	(.33)
PEDRO	(.50)	.57		0
ANTON	.75	.28		
FRANCO	.17	.18		
TONI	(0)	.43		
LEONARDO	(0)	.25		
CARMELA	(0)	(.67)		
LOLITA	(0)	(1)		
ANTONIO	0	0	0	
PASCUA	0	0	0	
ANGELINA	(0)	0		
ELISEO	(0)	0		
ROSEMARIE	(0)	0		

the labels used in this volume to clarify what corresponds to what, although we will not labour on the significance of the different functions of TOP.

The results shown in (6) support the Prominence Hypothesis because five learners use XP canonical word order exclusively (Antonio, Pascua, Angelina, Eliseo and Rosemarie), three learners (Janni, Agostino and Pinto) use noncanonical word order categorically in both declaratives and questions, and the remaining 17 learners use noncanonical word order at least sometimes. With respect to whether or not noncanonicity develops at the same time and in the same way in questions and declaratives, the table shows that XP canonical order is either co-present or entirely absent in declaratives and questions in all learners with the exception of four (Toni, Leonardo, Carmela and Lolita), who produce noncanonical word order in declaratives but not in questions. So, at least for these latter four learners, the table marks the emergence of noncanonicity for declaratives before questions. We note however that, after emergence, many more learners (11 of them) show categorical use of noncanonicity with questions, whereas only 4 (Janni, Agostino, Pintos, Maria S.) show categorical use with declaratives. We will take up this point in our discussion below.

In (7) we present the results of Pienemann's (1989; also 1987 and 1998) one-year longitudinal study of three ab-initio Australian university learners of German, which he conducted in order to compare their development to that of the naturalistic learners investigated in the ZISA cross-sectional study. As the grey area in the table shows, however, only one of his three learners was recorded for the whole year. Pienemann's data has been treated here in the same way as ZISA's have in (6).

(7) *Application of the SUBJ-V inversion rule (i.e., production of structures at the noncanonical word order stage) in questions and declaratives – ratio figures from Pienemann (1989: 68-70)*

LEARNER	ORIGINAL LABELS	OUR LABELS	WEEKS														
			1	3	5	7	9	11	13	15	17	19	21	23	25	27	
STEVEN	PP	TOP _{ADVERB}					0	0	(.5)	0	0	0	0	0	0	0	
	TOPI 2	TOP _{SUB. CLAUSE}															
	TOPI 1	TOP _{NP}										(0)				(1)	
	Q-PRO	FOC _{WH-}	.5	1	1	1	.75	1		1				1		1	
GUY	PP	TOP _{ADVERB}					0		0	0	0	0	0				
	TOPI 2	TOP _{SUB. CLAUSE}															
	TOPI 1	TOP _{NP}							0	0							
	Q-PRO	FOC _{WH-}				1	1	1	1	1	1	1	1				
VIVIEN	PP	TOP _{ADVERB}															
	TOPI 2	TOP _{SUB. CLAUSE}															
	TOPI 1	TOP _{NP}															
	Q-PRO	FOC _{WH-}	1	1			1	(0)									

Also the results from Pienemann's learners support the Prominence Hypothesis in so far as the learners reach the XP canonical word order stage (marked with a 0 ratio in the table) before or at the same time as the noncanonical word order stage (Vivien is an exception to this but, significantly, only in questions). Furthermore, noncanonicity emerges in questions long before declaratives in all three instructed learners, unlike in ZISA's cross-sectional study of naturalistic learners. In fact, all declaratives are produced without 'inversion' by the three learners, with the only exception of Steven, who produces one or two declaratives with 'inversion' in week 11 and then another few in the very last week, marked respectively with the ratios of .5 and 1.² Remarkably, also, noncanonicity is virtually categorical in questions, as the exceptions are rare and concern only the readings for weeks 3 and 11 for Steven and week 9 for Vivien.

4. The study

Our study on content questions in German L2 comprises a subset of the 16 learners who produce them out of the original 21 in Jansen's (2008) study, which dealt with declaratives only. All 16 learners are adult native speakers of English, ranging in age from 18 to 49 years; 8 are female and 8 male, and they are all students enrolled in a German course at an Australian tertiary institution. In terms of length of instruction, 10 learners (the first from left to right in (8) below) would be comparable to Pienemann's (1989) learners at week 13 because they had one semester of instruction. Only two of these learners have been to a German speaking country for one or two weeks as tourists. As for the remaining 6 learners, one of them (CR) had studied German for 3 semesters, and the other 5 (rightmost in the table) had studied the language for 5 semesters, including topical courses, such as literature and culture studies, delivered in German. Exceptionally, LD had completed a Science German course, had a more varied instructional background in German, including private tuition, and at the time the sample was taken was teaching a beginner German class in school. None of the learners had significant exposure to naturalistic spoken input, such as during conversation in a bilingual German-Australian home, or instruction in an immersion setting. However, many indicated that they had occasional exposure to naturalistic input outside of class, such as watching German films, reading in German or conversing with other nonnative speakers of German.

2 The general lack of raw numbers in the original studies (both ZISA's and Pienemann's) is rather unfortunate. In these two cases, however, we can desume the exact numbers of these occurrences from the original study.

To elicit the data, learners were asked to meet and have a conversation with a native speaker of German, whom they had never met before. The topic of the conversation was “getting to know one another”. The declared purpose was to “provide a speech sample for a research project on second language acquisition”. It was emphasized, to both the learners and the native speakers, that although the researcher was interested in a sample of the learner’s speech this should not encroach on spontaneity, the conversation should not be like an interview, and learners too were expected to ask questions. The conversations lasted about 45 minutes. They were audio recorded and transcribed for analysis.

In line with current PT analyses, and in contrast with the two studies considered in our literature review above (i.e., Clahsen, Meisel & Pienemann 1983 and Pienemann 1989), only sentences containing a lexical V are included in this study to the exclusion of copular sentences (e.g., *es ist gut* [‘it’s good’]) and presentatives (e.g., *es gibt nichts* [‘there is nothing’]). Excluded are also sentences with expletive SUBJ (e.g., *es regnet* [‘it rains’]) and those which are structurally ambiguous or otherwise insufficiently audible to be analysed; verbatim repetitions from the interlocutor, and typical (semi)formulaic questions frequently used in early classroom communication (e.g., *wie sagt man x?* [‘how does one say x?’]). Repeated identical sentences (i.e., those with the same lexicon as well as the same structure) are counted once. Thus the total number of sentences analysed here for the 16 learners is 1372. Among them, 85 are questions, and 1267 are declaratives. Among the latter 1059 display canonical word order, and 208 have a nonSUBJ element as TOP in first position. Among the questions, only 2 have in-situ FOC, and the rest display fronted FOC.

5. Results

Results of the analysis are illustrated in (8). The learners are referred to by a code and ordered according to the range of structures they actually produce. The left-most column shows the structures with FOC representing QPs, which may be arguments or ADJs; TOP represents fronted elements other than SUBJ in declaratives, which also may be arguments or ADJs. The numbers in the cells represent the frequency of the structure in the data. For instance, reading the learner YJ’s declarative results from bottom to top, we see that she produces 62 structures with canonical order, 4 XP_{TOP} with canonical order and 3 XP_{TOP} with noncanonical order. As for questions, she produces two (in-situ) canonical word order questions, and three XP_{FOC} with noncanonical order.

As (8) clearly shows, all learners produce canonical SVO/SOV sentences, and do so in greater numbers than all the other types of structures. Results for declaratives show that when learners add a topicalised element to their sentences all of

- b. CR 1972 habe ich nach England gegeben
 TOP_{ADJ} AUX SUBJ ADJ V
 1972 have I to England gone
 [in 1972 I went to England]

As can be appreciated from these examples, the range of structural exponents of TOP is wide, including: subordinate clauses as in (9a), NPs as in (9b), adverbs as in (10), and PPs as in (11a).

Results with questions differ significantly, and point to a strong role for FOC in bringing about noncanonical word order. First, only one learner (YJ) uses *in situ* FOC, as in (12), which belongs to the canonical word order stage. Interestingly, this learner also produces a FOC preceded by ADJ, as shown in (14), the only structure of this type in the whole data set.³

- (12) YJ du arbeitest da seit wieviele Jahren?
 SUBJ V ADJ FOC_{ADJ}
 you work there since how many years?
 [you have been working there for how many years?]

Second, whereas all 16 learners produce ungrammatical declaratives with TOP SVO structures, none of them produce questions with ungrammatical FOC SVO structures. Thus not only have they all reached the higher stage, but their word order is also applied categorically, including when they produce marginally acceptable structures, as MR in (13), as well as complex structures, as in (14), where FOC is preceded by TOP_{ADJ}.

- (13) MR wie lange hast Sie gelibt ins Australisch ?
 FOC_{ADJ} AUX SUBJ V ADJ
 how long have you lived in Australian
 [how long have you been living in Australia?]
- (14) YJ danach was willst du machen?
 TOP_{ADJ} FOC_{OBJ} MOD SUBJ V
 after that what want you do
 [after that what do you want to do?]

3 This complex structure, together with its accurate morphology, points to a more mature use of *in-situ* FOC questions, which is not unknown in native speakers speech (cf. Bettoni & Ginelli's comments with regard to Italian L2 *in-situ* questions, ch. 8, this book). Cf. also this same learner's complex structure in (14).

As with declaratives, also with interrogatives the range of structural exponents of the fronted element is wide: question phrases in (12)-(13), as well as question words in (14), may function as ADJ or argument GFs.

Admittedly, our figures are not always robust, as some of the learners (e.g., CP and RA) produce only one question with noncanonical word order. However, questions such as those in (15a-b) appear to be constructed online (i.e., nonformulaically) because a subsequent check confirmed that their Vs (i.e., *wohnen* ('live') and *gehen* ('go') respectively) are used also in other structural contexts.

- (15) a. CP wo wohnst du ?
 FOC_{ARG} V SUBJ
 where live you
 [where do you live?]
- b. RA wo gehen wir ?
 FOC_{OBL} V SUBJ
 where go we
 [where do we go? (he is asking himself aloud)]

6. Discussion

To recap, our cross-sectional study of 16 learners investigates the development of content questions in German L2 and compares the results both internally, with the development of declaratives, and externally, with relevant studies in German L2 such as ZISA's and Pienemann's. These two studies are different from ours not only because of their method of counting and presenting frequencies as ratios instead of raw figures, but also because they include copulas and presentatives. Significantly, however, a methodological approach proposed in this volume, but not yet applied empirically elsewhere (cf., however, the theoretical treatment of developmental syntax in English L2, § 3.1, ch. 2, this volume), is justified in the current study. We are referring to Bettoni & Di Biase's requirement for analytically sorting out questions from declaratives (cf. §§ 4.2.1 and 4.3, ch. 1, this volume) in order to clarify possible effects of different sentence types, such questions and declaratives, on timing of emergence and developmental patterns.

Our first research question asked whether or not the Prominence Hypothesis is supported for German L2. Results in all three studies, despite their differences in time, space and linguistic environment, are indeed compatible with the developmental schedule for German L2 based on the Prominence Hypothesis which we formulated comprehensively for both declaratives and questions in (5) above. In all

cases where noncanonical word order emerges, it does so either at the same time, or after canonical word order.

Our second research question asked whether noncanonicity develops at the same time and in the same way in declaratives and in questions. The answer appears to be negative for both variables. In terms of time of emergence, noncanonicity in questions emerges long before than in declaratives. In Pienemann's longitudinal study, in one of the learners (Steven), noncanonical word order ('inversion', in his terminology) emerges in questions in week 3, and in declaratives in week 11. In the other two learners, it emerges in questions in week 1 (Vivian) and week 7 (Guy), and does not emerge at all in declaratives. This earlier emergence in questions is confirmed in the current cross-sectional study, where all 16 learners use noncanonicity in questions, but seven of them fail to do so in declaratives. In contrast, in four out of the 25 ZISA learners noncanonicity emerges in declaratives before it does in questions, thus providing some counter-evidence to the contrast between questions and declaratives in terms of timing of emergence. In sum, noncanonicity in questions precedes noncanonicity in declaratives (with the four ZISA exceptions).

In terms of developmental patterns, noncanonicity in questions appears to be categorical, whereas in declaratives there is a persistent survival of XP with canonical word order even in those learners who do produce noncanonical word order comfortably. Thus, the difference between questions and declaratives is actually dramatic. Not only do all the 16 learners in our study produce postverbal SUBJ with fronted FOC in questions, but they do so categorically. This finds strong similarities in Pienemann's study, where the number of contexts is generally robust, particularly so in Guy's figures. The number of contexts in the ZISA study also provides some support for noncanonical word order in questions to emerge in a categorical fashion, certainly more readily in questions than in declaratives. The laborious progress of 'inversion' in the presence of fronted elements had also been noted by Clahsen (1984) in his longitudinal study of three ZISA learners, who observed that it is a difficult rule to learn. What has not been noticed so far, however, is that such slow and noncategorical progress may be confined to declaratives.

What remains to be explained, then, is why categorical noncanonicity in word order emerges earlier in questions, and lags so far behind in declaratives. What is specific to questions that can account for this? A couple of reasons may be put forward, at least for German L2. An often invoked possible reason is L1 transfer. The informants in the present study and in Pienemann's are from an English-speaking background, and English requires noncanonical word order in questions but not in declaratives. However, English requires *do*-support for questions unless the lexical V is already supported by an AUX or MOD V. In

German, on the other hand, such *do*-support is unknown and the lexical V is in sentence final position in the presence of AUX or MOD V. So, in spite of some similarities, the differences seem to be quite significant, and transfer, if any, would require extra computations in working memory. Testing the transfer hypothesis would in any case require triangulation with at least one other language with no postverbal SUBJ in constituent questions (cf. the DMLT Hypothesis in Pienemann, Di Biase, Kawaguchi & Håkansson 2005). A second reason may be that unlike TOP in declaratives, FOC in questions is obligatory (Mycock 2007), and therefore reliable in the input and close to one-to-one form-function mapping (cf. Andersen 1993; Ellis & Collins 2009, among others). Moreover, and most significantly, FOC in constituent questions is lexicalised transparently through a small, closed set of items, whose lexical specification includes focus, namely the *wh*-question words (Horvath 1986: 188). By contrast, TOP can be lexicalised in a virtually infinite number of ways. Therefore, the input for TOP is not only less frequent but also highly variable. Note also that in the very frequent default declaratives TOP usually coincides with SUBJ, which occupies the first slot in *c*-structure. This contrasts with the obligatory postverbal position of SUBJ in constituent questions (the only exception being when SUBJ itself is questioned). Postverbal SUBJ may thus be computed as the default position for SUBJ in content questions.

7. Conclusion

The results presented in this chapter, including those of the historical studies, support the Prominence Hypothesis. They document a robust continuity in SLA studies, but they also show that further analytical work based on sound theoretical principles can bring out hitherto unsuspected insights. Thus the separation of declaratives from questions enables the identification of significant differences in the acquisition patterns of noncanonical word orders. These differences are, on the balance of evidence, first that noncanonical orders emerge in questions before they do in declaratives, as confirmed in Pienemann's (1989) longitudinal study, and second that noncanonical word order in questions in German L2 is appropriated categorically by Jansen's (2008) learners, nearly so by Pienemann's, and tendentially also by ZISA's (1983). Hence, acquisition patterns are different: predominantly categorical in questions, and predominantly gradual in declaratives, where the canonical word order of the lower stages persists alongside the target-like noncanonical one of the higher stage even in significantly advanced learners.

A limitation on this conclusion is that the data for categoricity in questions is small in some learners given the naturalistic nature of the data collection method.

More robust data collected with focused tasks would better test these early results. If this chapter offers some possible explanations for the differences in the acquisition of noncanonical word order in questions versus declaratives, further research may reveal whether different GFs (argument or nonargument functions) and different elements (words, phrases or subordinate clauses) linked to FOC or TOP do or do not play a role in such remarkable differences, and may help establish a more unified and theoretically grounded explanation.

Thanks are due to Camilla Bettoni for her help in discussing and organising the issues dealt with in this chapter. We also wish to thank Wayan Arka for his comments on an earlier version of the chapter, as well as Gabriele Pallotti and two anonymous reviewers for their insightful comments.

*Exploring Processability Theory-based hypotheses
in the second language acquisition of a child
with autism spectrum disorder*

Tonya G. Agostini* and Catherine T. Best**

*University of Western Sydney, **University of Western Sydney and
Haskins Laboratories

1. Introduction

Delays and deficits in structural language and social communication are common features of autism spectrum disorder (ASD), but the exact nature of these problems is unclear. In particular, little is known about grammatical development in children with high functioning autism spectrum disorder (HFASD). An understanding of how language develops in this population may provide valuable insight into how underlying processing difficulties contribute to speech delays. The aim of this study is to explore whether a child with HFASD is able to acquire a second language. This study adopts Pienemann's (1998) classic Processability Theory (PT) approach in order to predict and measure how early inflectional morphology develops in this child when acquiring Italian L2, compared with typically developing children.

1.1. High functioning autism spectrum disorder

Autism spectrum disorder (ASD) is a lifelong neurodevelopmental disorder first described by Kanner (1943). The criteria for diagnosing ASD are based on impairments in two domains: (i) social communication and interactions, and (ii) restricted and repetitive behaviours and interests, often obsessive in nature (APA 2013). Language outcomes vary greatly (Kim, Paul, Tager-Flusberg & Lord 2014), but functional use of language prior to age 5-6 years is predictive of increased communication and independent living skills in adulthood (e.g., Lord & Venter 1992; Howlin, Goode, Hutton & Rutter 2004). Some distinct language features in children with ASD who acquire spoken language are: a prolonged period of echolalia, pronoun reversals, repetitive play with words, significantly better expressive language skills than receptive ones (which is the opposite of typically developing chil-

dren), difficulties in perceiving and attending to speech, and an unusual and idiosyncratic use of language (cf. Boucher 2012; Kim et al. 2014 for overviews). Pragmatic language deficits are universal in ASD and lead to difficulties in maintaining conversations. Social interaction skills are thus impaired.

High Functioning Autism Spectrum Disorder¹ (HFASD) is a label used by researchers and clinicians to refer to children diagnosed with ASD but without intellectual impairment, that is, with IQ above 70 (e.g., Mesibov, Shea & Adams 2001). Even among verbal children with HFASD there is considerable heterogeneity in language skills in the areas of semantics, morphology, syntax and phonology. At least two language-based groups have been identified: those with normal, or age-appropriate, linguistic skills (HFASD-N), as measured on standardised tests, and those with language impairments (HFASD-I), some of which are similar to those found in Specific Language Impairment (SLI: Kjelgaard & Tager-Flusberg 2001; Tager-Flusberg & Joseph 2003; Tek, Mesite, Fein & Naigles 2014). The participant in this study falls into the first HFASD-N group, with age-appropriate linguistic skills at the time of the study, despite experiencing language delays during the preschool years.

1.2. Grammatical development in autism spectrum disorder

Longitudinal studies during early stages of L1 acquisition could provide valuable insights into the underlying linguistic mechanisms that lead to delays and deficits experienced in ASD (Tager-Flusberg 2004). However, relatively few ASD studies have investigated grammatical development in either L1 (Boucher 2012; Kim et al. 2014) or L2 acquisition (Ohashi et al. 2012). This is due partly to the fact that a diagnosis occurs between 2 and 4 years of age (Coonrod & Stone 2005; Filipek

1 The term HFASD here replaces High Functioning Autism (HFA) to reflect the new DSM-5 (APA 2013) diagnostic terminology. HFA previously referred to a child diagnosed with *Autistic Disorder*, who has no intellectual impairment, IQ>70, as categorised under the criteria of the Diagnostic and Statistical Manual of Mental Disorders – Fourth edition (DSM-IV) (APA 1994). Importantly, a diagnosis of *autistic disorder* implied onset prior to 3 years of age *with initial language delays*. This was in contrast to other high functioning individuals, for example, those with *Asperger Syndrome* which implied an onset later than 3 years of age and *no early language delays*. However, because the DSM-5 (APA 2013) has since replaced its five distinct pervasive developmental disorders (i.e., autistic disorder, Asperger Syndrome, Rett's Disorder, child disintegrative disorder, pervasive developmental disorder not otherwise specified (PDD-NOS)) with one overarching category of *Autism Spectrum Disorder* (Volkmar, Reichow, Westphal & Mandell 2014) we will use HFASD, and then specify the participant's language history and status at the time of the study.

et al. 1999; Tager-Flusberg 2005), and partly to the sparsity of longitudinal studies since the introduction of robust and well-accepted diagnostic criteria.²

A recent study reports that children with ASD produce syntactically less complex utterances than typically developing children matched on non-verbal IQ, despite a similar vocabulary size (Eigsti, Bennetto & Dadlani 2007). With regard to early morphological development, studies on English-speaking children with ASD report conflicting results (Tager-Flusberg, Paul & Lord 2005). Some claim morphological development is deviant (e.g., Bartolucci, Pierce & Streiner 1980; Howlin 1984), others that it develops the same way as in typically developing children (e.g., Jarrold, Boucher & Russell 1997; Tager-Flusberg & Calkins 1990; Waterhouse & Fein 1982; Tek, Mesite, Fein & Naigles 2014). This controversy has partly been resolved with the identification of at least the two distinct groups among verbal children with ASD mentioned above, that is, those with apparently normal age-appropriate linguistic ability, and those with language impairments in phonological processing and grammatical morphology, some of which are similar to those found in children with SLI (Kjelgaard & Tager-Flusberg 2001; Roberts, Rice & Tager-Flusberg 2004; Tager-Flusberg & Joseph 2003; Tek et al. 2014).

Most recently, Tek et al. (2014) investigated grammatical development longitudinally in eight participants with a diagnosis of ASD and later categorised as highly verbal (ASD-HV) (age range = 2;2 to 3;1) in comparison to a group of 18 typically developing children (age range = 2;3 to 2;8), over a period of 12 months. They tracked increases of productive use of several aspects of basic grammatical abilities and found that the ASD-HV and TD groups both increased productive use of most of the same grammatical structures and at a similar pace. However, authors later noted that the ASD-HV group had been receiving an average of 14 hours per week of Applied Behaviour Analysis (ABA) therapy, which directly targets communication and speech production (Tek et al. 2014: 83). This greatly limits any conclusions that could be inferred from the results, and furthermore highlights the difficulties of obtaining true natural spontaneous L1 speech data for very young children diagnosed with ASD.

Researchers have started to monitor speech development of potentially high-risk (HR) infant siblings of older children with ASD. Because some of them may eventually receive an ASD diagnosis, this would potentially allow retrospective reconstruction of their language development (Tager-Flusberg 2005; Zwaigenbaum et al. 2005; Hudry et al. 2014). However, this approach is high-

2 That is, the diagnostic criteria provided in the DSM-IV (APA 1994) and ICD-10 (WHO 1993) and, more recently, the changes in diagnostic criteria and terminology published in the DSM-5 (APA 2013). The latter are hoped to be more sensitive to ASD behavioural symptoms in infants and toddlers under 3 years of age.

ly resource intensive and has a relatively low likelihood of payoff, as the recurrence of ASD within the same family is estimated at 18.7% (Ozonoff et al. 2011). In these studies, to date, researchers have only administered broad language measures at periodic time intervals of 6 months or greater (cf. Hudry et al. 2014 for an overview). In one such longitudinal study, authors report that even with a large number of HR infant siblings involved, they had modest sample sizes because of the necessary separation of the larger HR infant siblings group into smaller diagnosis-outcome-based subgroups, that is, those later diagnosed with ASD, those with atypical development but not ASD, and those with typical development (Hudry et al. 2014). In this study we take the novel alternative approach of examining the beginning acquisition of an L2 in a child with HFASD-N.

Studies of L2 and bilingual language acquisition in ASD are very few. In the absence of empirical studies, Jelinek, Toppelberg, Snow & Tager-Flusberg (1999) suggest that children with ASD should avoid L2 learning and bilingual education, reasoning that their pragmatic language deficits and difficulty in initiating and maintaining conversation will make them poor L2 learners. Yet there is evidence that at least some children with ASD can learn an L2 (e.g., Hambly & Fombonne 2012, 2014; Kanner 1971; Kay-Raining Bird, Lamond & Holden 2012; Ohashi et al. 2012; Petersen, Marinova-Todd & Mirenda 2012; Seung, Siddiqi & Elder 2006; Valicenti-McDermott et al. 2012). However, these L2 studies in ASD have been largely observational and/or exploratory and report a wide variety of outcomes (cf. Hambly & Fombonne 2014 for an overview). None of them report on grammatical development. Most useful would be information about how children manage the earliest building blocks of L2 acquisition. In this study we look at the development of productive use of L2 inflectional morphology in NPs. Mastering inflectional morphology requires the child to grasp that alternating a specific phonological subunit changes the meaning of a word.

1.3. *The targeted structures*

Italian L2 was chosen for this study because of its morphological richness and morphophonological suffix alternation patterns at NP level affecting Ns and their modifiers, such as adjectives (e.g., *gatt-o ner-o*, 'black.M.SG CAT.M.SG'). As a stem-based language, its nouns and modifiers take an obligatory suffix that indicates number and gender through a matrix of final vowel alternations, as represented in (1). For the analysis that will follow in § 3, it is worth noting that two of the three main classes of nouns and adjectives have *-i* in their plural form. For this reason, *-i* is considered the default plural ending, and is usually acquired before the more marked plural *-e* ending (Di Biase 2002; cf. also ch. 3, § 3.1, this volume). Unlike English, which adds the morpheme *-s* to noun in order to indicate plurality, Italian

number and gender features are fused into one vowel at the end of the noun and adjective. Furthermore, within a NP, the same values of both these features of the head N must be expressed on all nominal modifiers, such as determiners, demonstratives and adjectives. In other words, for our purposes here, learners must learn to compute the agreement within a NP, and mark adjectives in attributive position with the number and gender values of the head N. In this study we ignored agreement of the article, which according to Di Biase (2002) is not convincing evidence of the activation of phrasal procedure, for two different reasons: on the one hand, because of its frequency it is often formulaically learned with the N; on the other hand, because of the complexity of its form-function mapping, its completely accurate use is usually achieved very late in both L1 and L2 acquisition (Caselli, Leonard, Volterra & Compagnoli 1993).

(1) *Ending alternation in the main classes of Italian nouns and adjectives (after Vincent 2011, examples added)*

	SG	PL	GENDER	EXAMPLE
NOUNS	–o	–i	masc.	<i>cavallo/cavalli</i> [horse/horses]
	–a	–e	fem.	<i>scimmia/scimmie</i> [monkey/monkeys]
	–e	–i	masc. or fem.	<i>leone/leoni m.</i> [lion/lions] <i>tigre/tigri f.</i> [tiger/tigers]
ADJECTIVES	–o	–i	masc.	<i>giallo/gialli</i> [yellow]
	–a	–e	fem.	<i>rossa/rosse</i> [red]
	–e	–i	masc. or fem.	<i>verde/verdi</i> [green]

PT provides the theoretical framework (Pienemann 1998; ch. 1, this volume) for our examination of the acquisition of Italian number/gender morphology by our participant. This theory serves our purpose well for two main reasons. Firstly, it is psychologically plausible and formally testable (cf. ch. 1, this volume). Secondly, it has already been applied to Italian L2 (cf. ch. 3, § 3), and tested for both adults (Di Biase & Kawaguchi 2002) and typically developing children (Di Biase 2002), allowing a straightforward basis of comparison for our participant's pattern of morphophonological acquisition. PT proposes a universal hierarchy of specific procedural skills which allow for the development of L2 morphology according to a predictable and implicational order (cf. ch. 1, § 4.2.1, this volume). In this study we consider the first three procedures: the lemma access procedure, yielding single words and formulas used in an unanalysed way; the category procedure, yielding lexical form variation; and the phrasal procedure, computing agreement between

nouns and adjectives, as they apply to Italian L2 (cf. the developmental stages based on these procedures illustrated for Italian in § 3.1, ch. 3, this volume, and partly repeated here in (2)).

(2) *First developmental stages hypothesised for Italian L2 inflectional morphology in noun phrases (after Di Biase & Kawaguchi 2002)*

PROCEDURE	MORPHOLOGICAL OUTCOME/STAGE	STRUCTURE	EXAMPLE
PHRASAL PROCEDURE	NP morphology	agreement within NP	<i>un gatto nero / due gatti neri</i> [one black cat / two black cats]
CATEGORY PROCEDURE	lexical form variation	plural marking on N	<i>un gatto / due gatti</i> [one cat / two cats]
LEMMA ACCESS	single words; formulas	single words; formulas	<i>mi chiamo Piero [my name is Piero]</i> <i>un gatto / due gatto [one cat / two cat]</i>

1.4. Research questions

The broad objective of this study was to investigate whether a child with HFASD-N is able to learn Italian L2 grammar beyond the production of single words and formulaic expressions. Specifically, for moving from the lemma access stage to the first grammatical stage of form variation, we address two interrelated aspects of morphophonological alternation in the L2 that do not occur in the child's L1 (English). Development in Italian L2 requires that an English-speaking child learn two grammatical principles: the final vowel of an Italian N must be changed in order to mark plural number, and the specific vowel alternation used depends on the class of the N. This entails two novelties for an English-speaking child: first, number is marked by changing a final vowel, rather than by adding a consonantal suffix as in English; and secondly, Ns divide into classes that require different vowel alternations, unlike English, where purely phonotactic principles guide pronunciation of the final /-s/ for number-marking. Then, in order to progress to the phrasal procedure stage, gender and number values must be not only computed separately on head Ns and adjectives, but also unified between them. Because of the different N and adjective classes, our learner must be able to produce such NPs as *gatto nero* / *gatti neri* ('black cat / black cats') and *gallina nera* / *galline nere* ('black chicken / black chickens'). In this study the stimuli avoided N and adjective agreements with a combination of different vowel suffixes, such as *gatto verde* ('green cat'), *gallina verde* ('green chicken') for singular contexts, and *galline verdi* ('green chickens') for plural ones.

It is already known that early language development is delayed in the majority of children with ASD. Hence we address two possible sources of that delay: first, the social and communication deficits associated with ASD may decrease language input and interaction; secondly, language delays in ASD may be caused by underlying language processing difficulties/deficits. In the former case, if the primary cause of language delay is social, then a child with HFASD-N with no apparent cognitive and linguistic deficits should be able to acquire the means to mark the plural on Ns and compute the agreement within NP when taught under instructional conditions suited to the child's social and attentional limitations (cf. below). In the latter case, that is, if language delays are caused by language processing difficulties/deficits, we predict that, even under suitable instructional conditions, a child with HFASD-N learning Italian L2 should still demonstrate difficulties with acquiring the new morphophonological and grammatical principles of Italian. This would entail that our participant is slower than, or different from the typical children studied by Di Biase (2002), and that, in particular, the child with HFASD-N will first overextend the default singular forms to plural contexts, and then use the more common default vowel *-i* for marking plural Ns and adjectives for a longer period, and be delayed in learning to apply the rarer non-default *-e* ending.

2. The Study

2.1. *The participant*

The child participant, Chris (a pseudonym), is a six-year old boy from an Australian English L1 background. His parents became concerned about his language and social development prior to age 2;0 when he seemed to stop learning new words, had difficulty with social interactions, and frequently exhibited high levels of distress. Chris also failed to use gesture to communicate, displayed very little eye contact, and seemed to use language out of context, in an echolalic manner. At times he appeared deaf, not responding to his mother's voice nor turning his head towards the speaker when his name was called.

Chris was first diagnosed with ASD at age 2;4 by a developmental paediatrician, in accordance with DSM-IV criteria set (APA 1994). He immediately began weekly sessions of speech and language intervention. By age 5;1, his overall level of cognitive and language development and functioning was within the normal range, according to results in the Brigance (1991) K-1 screen, which assesses receptive language, expressive language and cognitive skills. As a result, Chris' diagnosis was reclassified to HFASD: his grammatical development was appropriate to his age, but his conversational skills were pragmatically impaired.

At the time of the study, Chris was aged 6;5 and his CELF-IV (Semel, Wiig & Secord 2003) results placed his speech production within normal range. However, he still presented with some receptive language difficulties relating to comprehension of pronouns.

Chris' paternal grandfather is a native Italian speaker, and both parents studied Italian L2 at university level. From age 3;0, the child had received some limited exposure to Italian L2 by means of Italian stories and cartoons. Moreover, he had attended an Italian story-time group for preschoolers. At the start of the study, he was thus able to count to 20, and label many animals, food items and colours in Italian, but he had acquired no grammar and was at the lemma access stage.

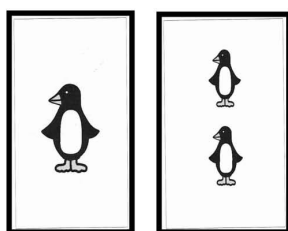
2.2. *The data*

To investigate the participant's grammatical development in Italian L2, a short-term longitudinal study was carried out. Over a 12-week period, Chris received 9 lessons of developmentally planned and communicative style teaching of Italian (cf. Di Biase 2002). Lessons lasted one hour each and were provided in a one-to-one teaching environment by an Italian native speaker. Data was collected five times, including a pre-test and post-test, then transcribed by the researcher, and checked against the video tape by an Italian native speaker. Rather than analysing Chris' linguistic production during the lessons, the acquisition of the targeted structures was monitored by means of specific tests. This is because, during the lessons, contexts contrasting the targeted number and gender forms for the same referent occurred rarely, making it difficult to determine whether emergence of these structures had occurred. Furthermore, during instruction sessions, it is difficult to determine how much the teacher's corrective feedback affects the child's performance. On the other hand, the tests were set at the very beginning of the lesson, and the teacher was instructed to refrain from providing any corrective feedback during the testing tasks. The five tests make up the corpus for this study. In addition to the pre-test in week 1 (t1) and delayed post-test in week 18 (t5), the other three tests occurred in lesson 4 /week 4 (t2), lesson 7 /week 7 (t3), and lesson 9/week 12 (t4). Each test lasted 5 to 10 minutes.

All five data collection sessions used picture-naming tasks. The stimuli were wordless picture cards, designed to elicit contrasting singular and plural Ns and adjectival NPs, and the child was asked to name the pictures in Italian, including their colour. An example of the stimuli is provided in (3). Most vocabulary items used for the tasks were familiar to the informant, but new items were also included to assess the child's skills in generalising number and gender feature markers. Other strategies designed to minimise formulaic responses (i.e., language 'chunks' or vocabulary items rote learnt in their plural or singular forms) included variation

of presentation (e.g., photographic stickers were used in one session, and hand-drawn pictures in the next); shifts in language domains (e.g., food items in one session, and farm animals in the next); change in the quantities of repeated plural probes (e.g., 47 penguins in one session, and 2 penguins in the next). The changes in the stimuli for each testing session helped ensure the learner was not producing unanalysed language chunks or routines he associated with specific stimuli, an important control to assure linguistic rather than associative rote knowledge, especially when testing children with ASD.

(3) *An example of the stimuli used for the picture-naming tasks*



In order to provide an indication of the child's progress over the period of the study, in (4) we present the numbers of Italian word types produced by Chris for each testing time, and the mean length of turns (MLT), that is, the average number of words per utterance he produced. When calculating MLT, the number of Italian tokens in a turn excludes: English or ambiguous words (e.g., *okay* or *no*, which can be English or Italian), repetitions, incomplete-, unclear- or echoic-items, fillers and hesitations. Also excluded are counting routines up to the final number followed by a N (e.g., in *un due tre giraffe*, only *tre giraffe* is considered). The number of turns refers to any turn in which Chris supplied Italian tokens. When more than one probe response was produced in a turn, only the longest response was counted for that turn. As the table shows, both the vocabulary size and MLT increased consistently from one session to the next. Cumulative vocabulary increased more than sixfold (from 24 word types in t1 to 165 in t5), and also MLT scores show good progress, almost doubling from 1.4 at t1 to 2.3 by t5.

(4) *Italian language production during testing times*

	T1	T2	T3	T4	T5
TOTAL N. OF LEXICAL TYPES	24	94	111	150	165
MEAN LENGTH OF TURN	1.4	1.8	1.9	2.0	2.3

3. Results

In accordance with PT's developmental and non-normative perspective, all occurrences of plural marking in the learner's system are counted, irrespective of accuracy. So, for instance, if the learner supplies a default plural *-i* suffix to mark a plural context, this counts as plural marking even if the native Italian system would use a different suffix for that context. Following Pienemann (1998: 132-140), the distributional analysis presented in (5) places plural *-i* and *-e* markings on Ns in obligatory contexts at the category procedure stage, and number and gender agreement between N and adjective(s) at the phrasal procedure stage.

(5) *Chris's progress from the pre-test to the (t1) delayed post-test (t5)*

STRUCTURE	T1	T2	T3	T4	T5
NOUN <i>-i</i> + TWO ADJECTIVES <i>-i</i>					-3
NOUN <i>-e</i> + ADJECTIVE <i>-e</i>				+3	-1
NOUN <i>-i</i> + ADJECTIVE <i>-i</i>		-2	+2-3>1	+1-2	+7-9>4
PLURAL <i>-i</i> ON ADJECTIVE		-1	+1-2	-1	
PLURAL <i>-e</i> ON NOUN	-3	+2 >1		+3	
PLURAL <i>-i</i> ON NOUN	+2-4>1	+11-1	+4-1	+7-1>1	

+	= supplied in obligatory context (correct)
>	= oversupplied in a singular context (error)
-	= not supplied in obligatory context (error)
empty cell	= no context

At t1, the pre-test provided the learner with the opportunity to produce 9 nominal forms in the plural, 6 of them with an *-i* suffix and 3 with an *-e* suffix. No contexts were provided for adjectival forms on their own or NPs requiring agreement. In the *-i* context, Chris produced two target-like forms. These however cannot be taken as sufficient evidence for productive use of number variation at the category procedure stage, as specified in (45), § 5, chapter 1 of this volume. This is because, although *cagnolini* ('little dog'), shown in (6), is correct, *pesci* ('fish') is used also in a singular context, so there is an instance of oversuppliance, as shown in (7).

- (6) t1 Chris uno due tre quattro cinque sei *cagnollino
 [one two three four five six little dog-MASC.SG]
 Teacher ah [filler]
 Chris cagnollini sei cagnollini
 [little dog-MASC.PL six little dog-MASC.PL]

- (7) t1 Chris *pesci (in response to a picture of one fish)
 [fish-MASC.PL]
 Teacher guarda . quanti? uno due tre quattro . quattro?
 [look . how many? one two three four . four?]
 Chris pesci
 [fish-MASC.PL]

In the remaining four plural *-i* contexts, the default singular forms were used: three times with an *-o* N (*cinque *gattino*, ‘five cat’; *due *cavallo*, ‘two horse’; *due *elefanto*, ‘two elephant’), and once with an *-e* N (**tigre*, ‘tigre’). Likewise, in all three plural *-e* contexts, Chris invariantly produced the singular suffix *-a* in both plural and singular contexts (e.g., *un giraffa* vs. *tre *giraffa*, ‘one giraffe’ vs. ‘three giraffe’). We can conclude that at the beginning of the instructional period, Chris was still at the lemma access stage.

At t2, after three weeks of Italian instruction, Chris showed remarkable progress. The test provided more varied contexts for producing plurals: 14 single-N contexts, 12 of them for *-i*, and two for *-e*; one single-adjective context in *-i*; and two phrasal contexts with both items ending in *-i*. All single Ns but one were correctly inflected for plural, whether the singular vs. plural alternation is *-o* vs. *-i* (*un coniglio* vs. *due conigli*, ‘one rabbit’ vs. ‘two rabbits’), *-e* vs. *-i* (*un maiale* vs. *tre maiali*, ‘one pig’ vs. ‘three pigs’), or *-a* vs. *-e* (*due mucche* vs. *una mucca*, ‘two cows’ vs. ‘one cow’). Furthermore, there was evidence that this target-like production extended to new lexical items. As shown in (8), when presented with the picture of an unknown referent (a camel), Chris first produced *un camel* (‘a camel’) with Italian phonology, which the teacher recasted with its proper singular *-o* suffix as *cammello*; then, with no prompts, albeit after a self-repair, Chris correctly produced *otto cammelli* (‘eight camels’).

- (8) t2 C ah well oh eh un camel (spoken with an Italian accent)
 [one-MASC.SG camel]
 T cammello
 [camel-MASC.SG]
 C cammello un cammello
 [camel-MASC.SG one-MASC.SG camel-MASC.SG]
 T è un cammello e questi?
 [it is one-MASC.SG camel-MASC.SG and this-MASC.PL?]
 C uno d oh I’m counting them uno due tre quattro cinque sei sette otto
 [one two three four five six seven eight]
 T otto?
 [eight?]
 C otto cammellia li ... cammelli
 [eight ... camels-MASC.PL]

On the other hand, at t2 Chris missed a plural suffix once, when presented with a picture of ‘many’ penguins. Despite their high number, he decided to count them all in Italian, but needed scaffolding from the teacher beyond number 20. When he finally counted the forty-seventh penguin independently, he produced *quarantasette* **panguino* (‘forty-seven panguin’, with a mispronunciation of a vowel in the stem and, more importantly, a singular ending). Furthermore, Chris also oversupplied plural *-e* on a N in a singular context (*un l g la* **galline* vs. *due la galline*, ‘one chicken’ vs ‘two chickens’), and left the unmarked default *-e* ending on the single adjective, producing **marrone* (‘brown’) instead of *marroni* with reference to *sette cammelli* (‘seven camels’) said by the teacher.

We can conclude that after only three weekly lessons, Chris had clearly moved one step up in the developmental path, from the lemma access stage to the category procedure stage. On the other hand, when provided with two contexts for a structure requiring the activation of the phrasal procedure, Chris failed to mark the agreement on the adjective, whether in the singular that ends in *-o* (*elefanti* **grigio*, ‘grey elephants’) or in *-e* (**marrone asini*, ‘brown donkeys’, with English word order), albeit in both cases the N was correctly marked as plural. As a matter of fact, it is worth reporting here that, before producing these two NPs, when first required to name animals and their colours, Chris seemed agitated and began with false starts and repetitions even if the adjective was *rosa*, which remains unchanged in all contexts. We could infer that at this stage the very production of a two word NP seemed to be a problem for Chris.

At t3, Chris continued to improve. He correctly produced four examples of plural *-i* marking on single Ns, using singular *-o* only once (*due* **pinguino*, ‘two penguin’). Interestingly, this occurred when Chris was prompted to produce this N with two adjectives (i.e., ‘two black and white penguins’), even though during the same task he had just produced the correct plural form on its own. Now that he could categorically mark Ns with confidence, Chris began to also mark adjectives at the category procedure stage, and produced a correct plural *-i* marker on a single adjective in one context (*bianchi* ‘white,’ referring to ‘penguins’) out of three. In the other two contexts the adjectives were used in their default singular form. A clear step forward at t3, however, was shown by agreement within NP, when Chris produced two target-like singular vs. plural agreement contrasts (*l’elefanti grigi* vs. *una l’elefante grigio*, ‘the grey elephants’ vs. ‘one grey elephant’; *un gatto nero* vs. *due gatti neri*, ‘one black cat’ vs. ‘two black cats’) out of a total of 5 contexts. In both of these cases, the plural adjectives require an *-i* suffix like their head Ns. In the remaining 3 contexts the adjectives bear default singular forms: *pinguini* **nero* (‘black.M.SG penguin.M.PL’); *due gatti* **arancione* (‘two orange.M.SG cat.M.PL’); and *pulcini* **giallo* (‘yellow.M.SG chick.M.PL’).

At t4, there are 7 examples of plural *-i* marking on Ns, five of which contrast with singular *-o* contexts, and two with singular *-a* contexts (e.g., *una torta*.F.SG vs.

due **torti*.F.PL, ‘one cake’ vs. ‘two cakes’). These latter two cases are not target-like, yet they were scored with a plus in (5) not only because Chris supplied a plural suffix but also because this overgeneralisation of *-i* provided further evidence that he was applying it productively. For the first time now, however, Chris also used *-e* for marking plural on Ns, and produced three examples of this suffix, albeit two of them were not target-like insofar as they contrast with singular *-o* contexts (e.g., *un fungo* vs. *tre* **funghie*, ‘one mushroom’ vs. ‘three mushrooms’). On the other hand, the third example contrasts with singular *-a* contexts accurately (*mela* vs. *mele*, ‘apple’ vs. ‘apples’). This step forward with the expansion of the *-e* production at category level was further confirmed at phrasal level. Whereas at t3 there were only two cases of *-i* + *-i* agreement, at this test session Chris produced three of *-e* + *-e* agreements, one of which is shown in (9).

- (9) t4 C una mela
[an apple-FEM.SG]
T bravo che colore?
[good what colour?]
C *rosso
[red-MASC.SG]
T ross?
[re?]
C I don't want . sso [(re)d]
T una mela rossa
[a red-FEM.SG apple-FEM.SG]
C una mela rossa
[a red-FEM.SG apple-FEM.SG]
T e queste?
[and these?]
C oh quattro mele rosse
[oh four red-FEM.PL apple-FEM.PL]

Notice, however, that here again Chris seemed to be reluctant to say longer phrases, and when pushed further his resistance became explicit: ‘I don’t want three words’. Nevertheless, we can conclude that his progress at t4 consisted mainly of a consolidation of the *-i* plural marker and an expansion of the *-e* marker, not only at category level but also within the phrase.

At t5, 6 weeks after the final 9th lesson, Chris’ progress continued insofar as he appeared to overcome his initial feeling of distress in producing longer phrases. He ventured in this direction in two different ways. First, he began to respond to task probes with three-word grammatical phrases consisting of a numeral, the head N and an adjective: this happened altogether 31 times out of 40 at t5. Of these 40 three-word phrases, 16 were in a plural context, and 7 out of them had the *-i*

marker on both elements, which he over-generalised several times in place of *-e*, as in *quattro *banani *gialli* ('four yellow bananas'). Secondly, for the first time, Chris tried to produce NPs with more than one adjective. Although he was unsuccessful in marking the plural for the second adjective (e.g., *pinguini neri *bianco*, 'white and black penguins'), the attempt is evidence of further ease in handling longer stretches of his L2 language.

In sum, our initial research question posed in § 1.4 can now be answered positively. First, we have shown convincingly that our participant with HFASD-N was able to acquire the morphophonology of an L2. Furthermore, we have illustrated that he acquired the plural marking on Ns and adjectives along the developmental path predicted by PT generally, and tested specifically among children learning the same Italian structures by Di Biase (2002). That is, Chris produced singular and plural contrasts first on single Ns and adjectives at the category procedure stage, and then also on Ns and adjectives at the phrasal procedure stage. Because no phrasal structure was produced before he could activate the category procedure, PT's implicational hierarchy is respected. Finally, data shows that Chris gradually learned to overcome his initial reluctance to produce longer strings in Italian, and eventually produced NPs that were both three word long and grammatically target-like.

With regard to the pace of his progress, a comparison between Chris and the typically developing children in Di Biase's (2002) study is not straightforward even though in both studies the instructional period targeted the same structures and lasted 12 weeks. On the one hand, Chris was 6;5 years old, had received minimal previous exposure to Italian, and was then taught for the study in a one-to-one situation. On the other hand, Di Biase's children were two-to-three years older, had already been learning Italian for three years (albeit without any progress beyond the lemma access stage), and were taught Italian all together in a large class. In (21) we show the delayed post-test results for Di Biase's children in the control and experimental groups regarding both single plural forms and phrasal agreement. Notice that in the Di Biase study both groups received focus-on-form type treatment similar to Chris', with the only difference that the experimental group received corrective feedback only on targeted structures. Figures in the table indicate that, while the overall improvement was remarkable, and particularly so in the experimental group, three children in the control group had not progressed beyond the lemma access stage, and that among those who had, the default *-i* marker was by far more accurate than the *-e* marker. Finally, even in the experimental group, who received the optimal treatment, most children (i.e., all but three: Chr, Kat and Lau) produce target-like plural agreement in less than half the required contexts. It is thus reasonable to conclude that Chris's progress was also well within the range of the average typically developing children. Further, he seemed to develop each new stage at a fairly rapid pace.

(21) *Delayed post-test production of plural forms and phrasal plural agreement in Di Biase's (2002) study*

CONTROL GROUP					EXPERIMENTAL GROUP				
	NOUNS		ADJECTIVES			NOUNS		ADJECTIVES	
	-I	-E	-I	-E		-I	-E	-I	-E
AMY	7	(2)	5		ADR	6	1	6	2
CAR	2		5		ALE	2	2	2	
CHR					ALL	3	4	7	5
COU					CHR	17		7	(1)
JOR	7	1	11	1	KAT	3	3	4	6
JOS	15		15		LAU	5	2	3	1
KIE	7	(1)	12		MAT	4	2	3	(2)
NIK	5	3	8		SAR	3	1	2	3
OLG		2			SOP	5	2	2	

figures represent plural *-i* and *-e* forms supplied in c. 20 total contexts for each child
 figures in brackets represent dubious forms

NOUN + ADJECTIVE		NOUN + ADJECTIVE	
AMY	4/11	ADR	4+1/12
CAR	3/12	ALE	2+1/9
CHR	0/10	ALL	3+1/9
COU	0/11	CHR	10/14
JOR	8+1/11	KAT	4+2/11
JOS	14/16	LAU	4+1/9
KIE	9+1/15	MAT	3+2/9
NIK	5/8	SAR	3/7
OLG	0/13	SOP	3/8

figures after the slash represent numbers of contexts
 simple figures before the slash represent *-i* agreement
 figures preceded by + represent *-e* agreement

4. Conclusion

This is a first study worldwide of grammatical development in L2 with a child with HFASD-N. Our results show that a child with HFASD-N can learn an L2 beyond the acquisition of unanalysed single words and formulas. In a 12 week period, as his Italian L2 utterances become longer, Chris developed from producing mostly invariant single words to being able first to mark number in Ns, and then to form agreement between head Ns and their modifiers within NPs. This development

follows the same route as that of typically developing children reported in Di Biase (2002), and at a comparable pace, indeed perhaps a faster one.

In § 1.4 we mentioned two possible sources of early language delay in ASD. The first hypothesis claims that social and communication deficits can lead to a lower amount of linguistic input, and hence a slower pace of acquisition; the second that an underlying language processing deficit/difficulty leads to difficulties in acquiring L2 grammar. The evidence gathered in this study clearly indicates that Chris does not display underlying linguistic deficits; rather, with instruction that is structured appropriately to address his social/communication difficulties, he was able to acquire Italian inflectional morphology in NPs by the normal route and at a comparable (or faster) speed to normal children. He even demonstrates the ability to mark plural with both the default plural *-i* suffix and the more marked non-default plural *-e* suffix. This discrimination was acquired earlier than expected and may reflect the *detail-focused* processing style of children with ASD (Happé & Frith 2006) and/or ability to *hyper-systemise* (Baron-Cohen 2009) novel L2 grammar.

Recent studies have shown that bilingualism in typically developing children improves cognitive functioning, particularly executive functions such as attention shifting and inhibitory control (Bialystok & Craik 2010). Since children with ASD are posited to have deficits in executive functions (Ozonoff, Pennington & Rogers 1991), SLA may prove to be of considerable cognitive benefit for these children, rather than posing problems for them as argued by Jellinek, Toppelberg, Snow & Tager-Flusberg (1999). Future research would need to involve large group studies of children with HFASD and typically developing controls. It would also need to enlarge its scope in many directions, including investigations not only of more structures dealt with by PT, but also of earlier stages of morphophonological development, such as *phonological constancy* (Best et al. 2009), that is, the ability to recognise spoken words even in the face of phonetic variability in the words' pronunciation across a variety of speakers and regional accents. Also, a finer-grained analysis of speech production at a *phonemic level* may bring to light important differences in the way children with HFASD perceive and produce novel speech and articulatory gestures.

The authors wish to thank Bruno Di Biase for guidance throughout the project, and Camilla Bettoni for help in organising the chapter; then, the college of Arts (University of Western Sydney), and Co.As.It. (Sydney), for financial assistance and providing an excellent native speaking teacher of Italian; and, finally, speech pathologist Debra-Ann Tanne for monitoring the progress of the participant for the duration of the study.

Connecting CALL and second language development: e-tandem learning of Japanese

Satomi Kawaguchi
University of Western Sydney

1. Introduction

More than 20 years have passed since Computer Assisted Language Learning (CALL) was introduced into L2 classrooms. The advantages of CALL in SLA are well documented (e.g., Thorne 2008). In particular, recent advancements in multimedia technology allow for ever more authentic communication exchanges in L2 interaction. Text message exchange known as chat offers great opportunities for the language learner to interact with a native speaker of a language instantly regardless of their physical distance. Among CALL activities, chat is one of the most researched areas (e.g., Donaldson & Kotter 1999; Schwienhorst 2002; Toyoda & Harrison 2002; Iwasaki & Oliver 2003; Kotter 2003; O'Rourke 2005). Most studies in the area of chat are conducted within an interactionist approach. The Interaction Hypothesis (Long 1996) claims that interaction plays an important role in L2 acquisition, which is promoted through negotiation of meaning such as clarification requests and confirmation checks. In L2 learning, language input (Krashen 1985; Long 1996) and learner output (Swain 1985, 1995) play crucial roles. CALL can enhance both. In particular, online chat creates authentic contexts for interaction where the language learner has opportunities to receive meaningful input and output. Furthermore, negotiation of meaning and feedback on his/her output from the native speaker may promote 'noticing' the gap (Schmidt & Frota 1986) between the current state of the interlanguage and the target language.

The type of CALL activity chosen in this chapter is electronic tandem (e-tandem for short, Cziko 2004), an activity in which participants engage in tele-collaboration (Ware & Cañado 2007) via text-based Synchronous Computer Mediated Communication (SCMC). In L2 e-tandem learning, a group of L2 students engages in a learning interaction with another group of students who are native speakers of that language. These, in turn, are also learners of a second language, which is the native language of the first group. So, each group is, alter-

natively, learning from, or teaching the other group (cf. Lewis & Walker 2003 for detailed explanations of tandem language learning). In recent years, research on e-tandem has grown in the field of SLA. Studies on e-tandem learning show how learners can use negotiation opportunities qualitatively and quantitatively when communication problems occur (Iwasaki & Oliver 2003; Sotillo 2005; Lee 2006; Ware & O'Dowd 2008; Bower & Kawaguchi 2011). Other CALL-related studies deal with technological design, students' evaluation or perception and cultural issues. So far few studies on e-tandem have examined the effectiveness of CALL activities on L2 development (e.g., Bower & Kawaguchi 2011; Iwasaki & Oliver 2003). For learners, teachers and researchers it is important to be aware of language development when experimenting and practicing CALL in order to make it more meaningful.

This chapter presents an early attempt to link CALL and PT. I will use as an example e-tandem learning via chat between learners of Japanese L2 in Australia and learners of English L2 in Japan, with the focus on Japanese L2. I will analyse chat logs according to PT's developmental stages, and measure the learners' L2 production as the text chatting activities progress over a two-month period. Furthermore, I will attempt to link CALL to my main research line in this field, which aims at extending the application of PT's Steadiness Hypothesis across learning modalities, and test whether the production of text chat language conforms to the PT schedules formulated for orality. The Steadiness Hypothesis was originally proposed by Pienemann (1998: 273-297) when he convincingly showed that acquisitional sequences are not affected by different communicative tasks. So far only Håkansson & Norrby (2007) and Rahkonen & Håkansson (2008) have investigated steadiness in a modality other than speech, namely writing. Rahkonen & Håkansson's (2008) cross-sectional study on L2 Swedish writing looks at formal and semiformal writing. Their results show that acquisition patterns in the two types of writing present both similarities to, and differences from PT sequence for speech. That is, using the old PT terminology of these two researchers, the sequence in speech is 'tense < VP agr < INV'; in the semiformal writing corpus 'tense = VP agr < INV'; and in the formal writing corpus 'tense = VP agr = INV'. Thus, acquisitional patterns may differ significantly according to whether corpora include semiformal writing closer to speech or formal writing. So far, no study has been undertaken on SCMC within the framework of PT. It is interesting to examine whether the L2 text production via SCMC shows a close relationship with PT stages in speech, since it occurs in real-time interaction and is thus closer to online speech production than either formal or semiformal writing.

In order to bridge the gap in research the following research questions are addressed in this chapter:

- Do tandem learning activities conducted through chat promote L2 production? More specifically, does the number of word types produced by L2 speakers increase during e-tandem learning activities?
- Does online production in text messaging (written production) show development as measured by PT developmental stages?
- Is development across modalities steady in terms of the trajectory of morphological and syntactic structures defined by PT's Steadiness Hypothesis for oral production?

2. Synchronous CMC and SLA research

Communication technology enables us to interact with others in various ways. For example, e-mail exchange is an asynchronous way of Computer-Mediated Communication (CMC), whereas text chat is synchronous, the “most interactive end of the CMC spectrum” (Paramskas 1999: 17). As such, online text messaging shares many characteristics with face-to-face conversation. Participants have far less time to edit a message than in other types of CMC such as composition in blog or e-mail (Levy & Stockwell 2006). Because of real time interaction, they can negotiate meaning in a way similar to face-to-face conversation (Blake 2000; Toyoda & Harrison 2002; Iwasaki & Oliver 2003). Furthermore, the text-based medium can lead learners to noticing more problematic L2 language than face-to-face communication (Lai & Zhao 2006), and may increase students' attention to linguistic form (Warschauer 1996).

Payne & Whitney (2002) claim that, in L2 learning, chat may achieve better outcomes than face-to-face activities because the L2 learner benefits from slower language processing while at the same time having to process utterances largely ‘on the fly’, that is, with little or no advance planning, as is the case of face-to-face verbal communication. These authors demonstrate that a group of learners with blended activities (both face-to-face and SCMC) improve their oral proficiency better than an equivalent group with face-to-face learning only, when measured by Oral Production Interview ratings on comprehensibility, fluency, vocabulary use, grammar, and pronunciation.

Payne & Whitney (2002) also show that SCMC can develop the same cognitive mechanisms underlying spontaneous oral communication, and thus facilitate the acquisition of L2 speaking skills. Part of their claim is based on Levelt's (1989) Speech Model (cf. § 2.1, ch. 1, this volume) in relation to the role of working memory for language processing. Working memory has a limited capacity, and can thus attend only to a limited amount of information immediately in real-time. There is a trade-off between its information processing role and its storage role in order to cope with the incremental nature of language production/comprehension

(Kahnemann 1973; Carpenter & Just 1989). This means that if a speaker processes information more efficiently, more working memory space becomes available to devote to other components of language processing and storing. Beginning L2 learners need to place attentional resources on various steps of language processing. For example, they need to search for appropriate words from their mental lexicon and place them within an appropriate syntactic frame with appropriate functional assignment. Furthermore, they need to determine the correct morphosyntactic forms and select corresponding phonological units. If these processes are not automatic, a burden is placed on the phonological processing (Baddeley 2007). In L1, lexical access and articulation are largely automatic, and the speaker needs to pay attention only to conceptualization and careful delivery. On the other hand, at the earlier stages of L2 learning, most processing components are controlled, resulting in speech with longer and more frequent pauses, etc. (Poulisse 1999). PT also uses Levelt's (1989) Model and claims that L2 acquisition is constrained by the learner's current state of procedural skills. As particular procedural skill components are automatised, more attentional resources become available to the L2 speaker, who can thus achieve further language processing, such as more fluent speech or/and construction of higher stage-structures.

What may then be the advantages of using text chat communication in SLA? First, the speed of information exchange is a little slower than in speaking simply because one cannot write as fast as one speaks. This gives L2 learners opportunities to process L2 messages at a slower pace while going through a similar language processing as in face-to-face communication. Secondly, the availability of previous messages (the visual co-text) can help learners to reduce the amount of information to be stored in working memory. Therefore, they can free up more attentional resources for the L2 lexicon and structures while maintaining pace in interaction. Thus learners may gain greater benefit from slower language processing while going through similar processing as may occur in face-to-face communication involving conceptualization, formulation, articulation (typing) as well as comprehension of the interlocutor's utterance. Thirdly, Payne & Whitney (2002) believe that chat communication may provide opportunities for learners to experience a sort of "conversation simulator" which may be effective especially for less confident, shyer or linguistically weaker students, who tend not to take full advantage of interaction with teacher or fellow students in face-to-face settings. Furthermore, the text-based medium may increase students' attention to linguistic form (Warschauer 1996) due to the slower information exchange.

Chat logs obtained from a tandem project can therefore offer an interesting opportunity to investigate whether L2 development follows PT stages also in this communication mode. In order to measure learners' oral proficiency, Payne & Whitney (2002) use Oral Production Interview rating on comprehensibility, flu-

ency, vocabulary usage, syntax, and pronunciation with two examiners rating on a 50 point scale. Although interrater reliability was high (0.86 for the pre-test and 0.94 for the post-test), it might be fruitful to assess L2 development using a more objective measurement such as the PT stages.

3. The e-tandem project

This section describes a tandem learning project¹ using chat via instant messaging between language classes at the University of Western Sydney in Australia and at the Kanda University of International Study in Japan. This project was organised as an out-of-class L2 activity at each institution aiming at collaborative learning in a more authentic context.

The participants are 21 second-year students of L2 Japanese in Australia, and 21 first-year students of L2 English in Japan.² The two groups are compatible, albeit not perfectly matched in several ways. First, the levels of L2 English in Japan are generally higher than those of L2 Japanese in Australia, because Japanese students learn English for six years at secondary school as a compulsory subject before starting university, whereas most Australian students start learning Japanese at university. Secondly, all students of English in Japan are native speakers of Japanese, whereas the students of Japanese in Australia are a mixed group, reflecting the multicultural nature of Australian society. So, according to the answers to a questionnaire enquiring about their background, 10 students are foreigners, four identified themselves as immigrants of ethnic background, and the rest as Australians. In order to enhance compatibility, when the project started all learners wrote a short self-introduction so that their teachers could match the tandem pairs according to mutual interests.

The project includes three chat sessions, distributed over two months, each lasting at least 30 minutes in each of the two languages. The first session was conducted during class time to ensure that everything worked; the second and third sessions were then organized by tandem pairs autonomously.

Before each chat session, students were given a broad conversation topic. For session 1 this was more oriented towards the here-and-now, and dedicated to self-

1 This study was supported by the University of Western Sydney, LTAP (Learning & Teaching Action Plan).

2 Two second-year students at Kanda University of International Study and four third-year students at University of Western Sydney also volunteered on several occasions in order to replace students who were unable to attend, and thus match the number of participants.

introduction and one's family, then it gradually moved on to more challenging areas: session 2 was about university life, and session 3 about each other's culture and related controversial issues. Participants were required to do the following:

- organise the date and the time of the two subsequent sessions via e-mail, and agree on the topics for conversation;
- check and study vocabulary which may be useful for conversing on the agreed topics;
- log in and participate in the chat at the designated time, and converse on the chosen topics half the time in English and half in Japanese at each session;
- submit to their teachers their observations and chat logs after each session;
- study the tandem partners' L2 production in the chat log, and send them some corrections and suggestions via e-mail after each chat.

In § 4 below I will report on the analysis of L2 Japanese data only. Because not all participants attended all chat sessions, some students had a three-way chat or different partners across sessions. I thus selected five learners of Japanese who participated in all three sessions with the same tandem partner.

As mentioned above, the e-tandem sessions were an out-of-class activity organised over two months as part of the Japanese L2 course. In the face-to-face classroom, new lexicon and the following grammatical structures were introduced weekly during this period:

- potential forms of the verb;
- conditional clause;
- various interrogative sentences (asking about subject/object/adjunct);
- auxiliary verbs;
- plain forms of adjective, copula and verb;
- sentential modification for noun.

Also, learners might have learned new words and structures from their partners during the sessions. However, learning a new structure as knowledge is different from using it in time-constrained conditions. In order to produce the structure online, processing efficiency needs to be attained. In my analysis here, although development will be measured as the three sessions progress, the main focus is on language use rather than acquisition. Indeed, the e-tandem project is not designed to test the learner's acquisition of particular morphological structure (e.g., verbal inflections) and syntactic structure (e.g., passive voice) in order to identify his/her PT stages. Instead, what is clearly seen from the chat log is each L2 learner's language use (such as selection of lexicon and syntactic structures) in real time interaction.

In order to exemplify some of the characteristics of text chat such as the pattern of turn taking, an excerpt from session 3 by Lee and his Japanese partner Mayumi is reproduced in the Appendix. Text chat and face-to-face conversation share many similarities including negotiation of meaning, such as the confirmation checks in turn 2 and 3, and the clarification requests in turns 17-18. Also, since the conversation is between a learner of Japanese and a native speaker, code-switching can occur (e.g., turns 2 and 5). On the other hand, there are some differences between this modality and face-to-face conversations. For example, the pattern of turn taking is different. In chat, turns are counted whenever a typist posts the text message by pressing the “enter” key. Thus, unlike in face-to-face conversation, participants may have different turn-taking opportunities, because a fast typist may post another text message (e.g., turns 7, 9 and 16) before his/her partner responds to the previous one, which may result in uneven distribution of turns across interactants. Sometimes a typist may accidentally post a turn before he/she finishes a sentence and continues in the next turn (e.g., turns 25-26). Conversational sequence can be interrupted by some unexpected turns (e.g., turn 20) due to the delay/gap between the time of posting a message and the time of receiving the response. Furthermore, sometimes a participant may move on to a new topic before a negotiation is resolved (e.g., turn 20).

4. Language Development

4.1. Chat production

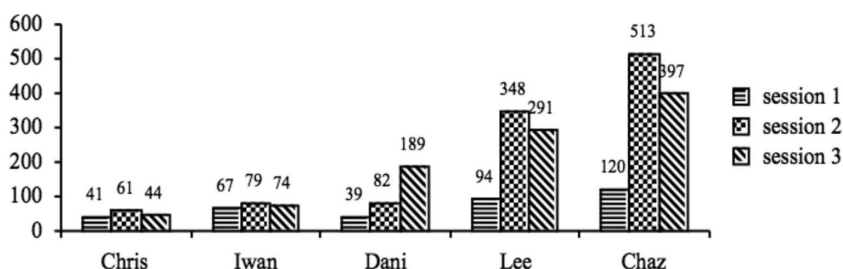
The numbers of turns produced in each chat session by our five learners are shown in (1). In the table “total” indicates the turns produced by both interlocutors, namely the native speakers of Japanese (NSs) and the L2 learner (NNSs); “NNS” indicates the turns produced by the L2 learner only; and the numbers in brackets indicate the turns produced in English.³ So, for example, Chaz’s session 2 comprises 109 turns, with Chaz himself posting 66 of them, three of which are written in English. Looking at all five learners, we can see that the text messaging activity seems successful, in so far as the number of turns sent in L2 Japanese increases for all of them after the first break-in session, although more so in session 2 than in session 3.

3 Codeswitches were encouraged by the topic, especially in session 3, when controversial issues in each other’s culture were discussed. For example, two pairs talked about “whaling” in Japanese. One of these pairs sometimes used English words/sentences when communication problems occurred.

(1) Number of turns by learner and session

	CHRIS		IWAN		DANI		LEE		CHAZ	
	Tot.	NNS	Tot.	NNS	Tot.	NNS	Tot.	NNS	Tot.	NNS
SESS. 1	24 (0)	13 (0)	25 (0)	15 (0)	18 (1)	10 (0)	54 (0)	26 (0)	26 (0)	16 (0)
SESS. 2	32 (3)	15 (1)	51 (2)	24 (2)	25 (0)	14 (0)	188 (9)	100 (9)	104 (5)	63 (3)
SESS. 3	17 (0)	9 (0)	59 (10)	29 (5)	31 (0)	17 (0)	152 (3)	70 (2)	89 (2)	47 (1)
TOTAL	73 (3)	37 (1)	135 (12)	68 (7)	74 (1)	41 (0)	394 (12)	196 (11)	219 (7)	126 (4)

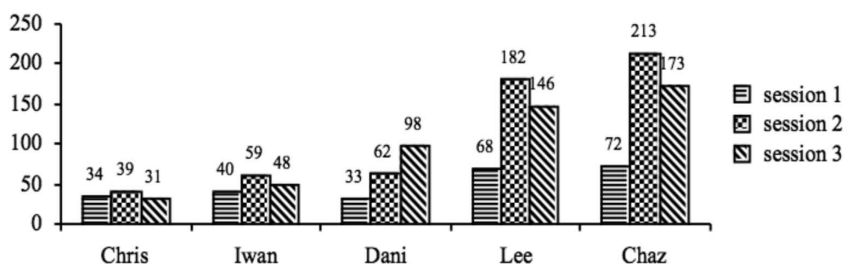
Because a turn may consist of one word or a few sentences, a more precise measure of the amount of language used is the word count in terms of tokens, as shown in the graph in (2). Here we can see that all five learners increase their tokens from session 1 to session 2, but then four out of five decrease them in session 3, as they do with their turns. Dani is the only learner who increases the number of both turns and words in the third session. The reason for smaller production in this session can be found in the topic of the conversation: whereas the topic for session 2 is “university life” and could be handled more descriptively, the topic for session 3 engages the learners on controversial issues relating to each other’s culture. Its argumentative nature may have slowed down the conversation quantitatively as learners would have required more time for conceptualising the message. In the graph we also notice that, if the curve seems similar across all learners except Dani, there is great variation among them in the number of actual tokens produced. So, in session 2, for example, Chris posts 61 word tokens and Chaz 513.

(2) Number of words (tokens) by learner and session

4.2. *The lexicon*

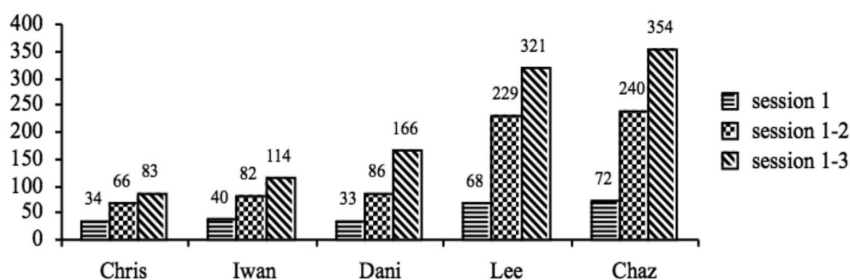
Using the Key Word In Context concordance programme (KWIC, created by Nihon University and available online), I calculated the size of the lexicon used during the chat sessions in terms of types (i.e., different words). This is illustrated in the two graphs in (3)-(4). The graph in (3) shows number of types used by each learner in each session. Considering the results of the token analysis in the graph in (2), it is not surprising to see that as the tokens vary, so do the types; that is, numbers for types increase for all students between session 1 and 2, but keep increasing only for Dani between sessions 2 and 3.

(3) *Number of words (types) by learner and session*



Because the topic of conversation changed in each chat session, and learners were required to study vocabulary relating to the agreed conversational topics prior to the chat session, we expected learners to produce some different lexical items every time. In order to measure this, the graph in (4) shows the cumulative number of word types without including those already used in the previous session; that is, for each learner, the types of session 1 are added to those of session 2, minus those already used in session 1; and the types of sessions 1 and 2 are added up to those

(4) *Cumulative number of word types by learner and session*



of session 3, minus those already used in sessions 1 and 2. Thus the figures in this graph indicate the actual size of the lexicon used over the three sessions, and show that the number of types steadily increased at every session for all learners. Comparing learners, however, we can see great variation among them, as we have already noticed with regards to their word tokens. For example, the figure for Chris in all three sessions is 83, for Dani 166, and for Chaz 354.

4.3. Morphology and syntax

The grammatical development of the five learners from session 1 to session 3 is illustrated in (5). For the description of PT stages for Japanese L2 morphology and syntax, I refer to chapter 4, §§ 2-3, this volume. However, as a matter of convenience, the first stage of lemma access is omitted here, because all learners already safely reached it. In the analysis, as mentioned in chapter 1, § 5, this volume, lexical or formal variation is used as the acquisition criterion for morphology, whereas for syntax one occurrence of the relevant structure is considered sufficient.

(5) The learners' progress over the three chat sessions

STAGES	STRUCTURES	CHRIS			IWAN			DANI			LEE			CHAZ		
		SESSIONS 1 2 3	SESSIONS 1 2 3	SESSIONS 1 2 3	SESSIONS 1 2 3	SESSIONS 1 2 3	SESSIONS 1 2 3	SESSIONS 1 2 3	SESSIONS 1 2 3	SESSIONS 1 2 3	SESSIONS 1 2 3					
M O R P H O L O G Y																
S-PROCEDURE	noncan. case marking							(1) 1			-2			(1)		
PHRASAL PROCED.	V-te V							2 3		1	19 20			4 15 2		
CATEGORY PROCED.	V inflection	2 4 2			3 2 3			2 10 11		23	22 11			4 26 27		
S Y N T A X - P R O M I N E N C E H Y P O T H E S I S																
NONCANON. ORDER	TOP _{OBJ} SV										1 4 1			3 2		
XP CANON. ORDER	TOP _{ADJ} SOV		2 1		2 -1 2				3		1 11 8			3 1		
CANON. ORDER	SOV	1 5 1			1 1 1			2 8 10		5	18 8			9 16 9		
S Y N T A X - L E X I C A L M A P P I N G H Y P O T H E S I S																
NONDEF. MAPPING	passives/benef.							1 1						1		
DEFAULT MAPPING & ADDL. ARG.	ag./exp. on SUBJ, pat./th. on OBJ, & goal/ben. on OBL		4 1		1 2			2 3		3	5 3			4 6 6		
DEFAULT MAPPING	ag./exp. on SUBJ and pat./th. on OBJ	1 5 1			1 1 1			2 8 10		3	14-1 14			9 16 9		

As (5) shows, Chris's numbers for both morphological and syntactic structures are small, because most of his turns are single words or fragments of phrases. However, he shows some progress during the three sessions. Morphologically, he stays at the category procedure stage in all three sessions, but he produces a variety of forms: polite (affirmative), negative and question forms in session 1, then adding the plain form in sessions 2 and 3. Syntactically, he moves one stage up with both word order and lexical mapping in session 2, then proceeds no further.

Also Iwan produces few sentences in the three sessions as he often uses fixed expressions. In terms of morphology, he produces only structures requiring category procedure (i.e., V inflection) in all three sessions. In terms of syntax involving word order, he is at the XP canonical word order stage at session 1 and does not progress further by session 3. In terms of syntax involving mapping, he is at the default mapping stage at session 1, and progresses to default mapping and to the additional argument stage at session 2. We can conclude that his progress is limited.

Dani's progress, on the other hand, is remarkable in several ways. Although, like Chris, he starts off in session 1 with only low numbers for the earliest stage in both morphology and syntax, he then moves up two stages in both the morphological and the lexical mapping schedules. Furthermore, morphologically both the numbers and range of structures increase substantially in sessions 2 and 3. That is, he produces 2 V inflections (one negative and the other desiderative) in session 1, which belong to the category procedure stage. Then he adds more V inflections, including past tense and plain forms, as well as *V-te* V structure expressing progressive aspect in session 2, which belongs to the phrasal procedure stage. He also produces noncanonical case marking once in sessions 2 and 3. Finally, although only one token of noncanonical case marking belonging to the S-procedure stage is produced, it can be assumed that he has acquired it by session 3 because he has already produced this structure with different lexical items in session 2. In terms of syntax, dramatic achievement is observed. Despite producing only sentences with canonical word order and default mapping in session 1, Dani constructs sentences with benefactive and passive V's in sessions 2 and 3.⁴ He also produces XP plus canonical order in session 3. In addition, in each of the last two sessions he actively uses further new structures learned in class: conditional clause in session 2, and sentential noun modification in session 3. Compared to Chris, Dani then progresses enormous-

4 Dani's production of passive sentence is as follows: 日本で 何頭クジラが 毎年 殺されるか (Nihon-de nantoo kujira-ga maitoshi korosareru ka) "In Japan how many whales are killed every year?"

ly over the two months, shooting ahead by two stages both morphologically and syntactically, and using twice the number of lexical types. He is also the only learner who never codeswitches to English.

Lee is at the category procedure stage in morphology in session 1 as he produces V inflection 23 times. In this session, he also uses one *V-te V* structure, which however is insufficient to place him at the phrasal procedure stage yet. In session 2, remarkably, he produces the *V-te V* and other combinatorial structures 19 times, and thus safely reaches the phrasal procedure stage. In session 3, he produces morphological structures requiring the category procedure 11 times and the phrasal procedure 20 times, and attempts noncanonical case marking (i.e., S-procedure) twice, but both these cases turn out to be wrong. Thus Lee's morphological progression is from the category procedure stage in session 1 to the phrasal procedure stage in sessions 2 and 3. As for syntax, in session 1 Lee is already at the noncanonical word order stage when, besides structures of the lower stages, he also produces one TOP_{OBJ} SV structure. In sessions 2 and 3 this rule is consolidated. In terms of lexical mapping, he is at the intermediate stage of default mapping and additional argument in session 1, and remains there in sessions 2 and 3.

Chaz also improves over the sessions. In session 1 he starts off with robust figures for both morphology at the phrasal procedure stage and syntax at the intermediate lexical mapping stage, and he is the learner who progresses furthest in syntax in the subsequent two sessions. In morphology he then seems to reach the S-procedure stage with just one tentative production of noncanonical case marking in session 2, shown in brackets in (5). In syntax, on the other hand, he comfortably reaches the noncanonical word order stage by producing several OBJ topicalisations. Furthermore, he produces one benefactive structure in session 2.

In sum, the analysis of grammatical development over the three sessions indicates that the answers to my research questions formulated in § 1 are all positive. The first question asks whether the cumulative number of words produced by L2 speakers increases during e-tandem. As the sessions progress, all learners' use more language, in terms of both turn numbers and word tokens. Also word types increase dramatically, although with significant individual differences (almost 1:4.5). The same tendency is apparent in the development of morphology and syntax. Two learners (Dani and Lee) out of five progress to new morphological stages. With regard to syntax, three learners progress on word order (Chris, Dani and Chaz), and four extend their mapping (the exception is Lee).

The second question asks whether developmental sequences in text chat follow the trajectory defined by PT for oral production. They certainly do, as these results demonstrate, with no learner skipping stages. As for the third question, whether PT stages hold across modalities, the answer is positive and detailed results also show that chat may produce results closer to speech than formal and informal writing. Thus, e-tandem chat is shown to be an effective, guided L2 learning activ-

ity which can increase L2 production. With regard to the large individual differences in students’ learning outcomes, at this stage we cannot say whether they are due to a different response to this set of activities.

5. Conclusion

CALL is now a natural part of L2 learning (Chambers & Bax 2006). This requires new roles for the language teacher, who must no longer just teach “knowledge” of the L2 in the classroom, but also design an L2 learning environment which promotes and supports meaningful L2 activities (Thomson 2007). CALL can offer these activities for out-of-class practice, thus offering new opportunities to achieve better L2 outcomes. However, there may be a pitfall with CALL if SLA theories are not taken into account. For example, CALL activities may seem to suite learners’ current needs and lifestyles, but they run the risk of being just an enjoyable experience and produce little L2 progress if teachers do not monitor the L2 learners progress using a reliable measurement such as PT stages.

This chapter shows that PT has a positive potential to contribute to the field and is capable to connect CALL and SLA. Further, the Steadiness Hypothesis is confirmed not only across tasks but also across modalities. This suggests a great potential for an online extended version of PT’s Rapid Profile (cf., e.g., Pienemann & Mackey 1993; Kessler 2007; Pienemann & Kessler 2010) in monitoring L2 grammatical development with CALL (especially text messaging) by teachers or learners themselves.

Appendix

A chat log excerpt from session 3 with Lee (L, learner) and Mayumi (M, native speaker)

TURN	CHAT LOG [ROMANISED TRANSLITERATION]	ENGLISH TRANSLATION
1 L	オーストラリアの海でくじら进行くことはいほうです くじらを殺すことは違法です [Oosutoraria-no umi-de kugira-o korosu koto-wa ihoo desu kujira-o korosu koto-wa ihoo desu]	It is illegal to kill whales in Australian waters. It is illegal to kill whales.
2 M	かいいき=a sea area をまもれってこと？ [kaiiki=a sea area o mamorette koto?]	Do you mean we [should] abide by a sea area?
3 M	そうなのですか？ [soo nano desu ka?]	Is that so?
4 L	そう [soo]	Yes.

5	L	international waters, japanese waters etc	International waters, Japanese waters, etc.
6	M	たぶん 日本のくじらをとってるひとは そのことを知りません [tabun nihon-no kujira-o totteru hito-wa sono koto-o shirimasen]	Maybe those who are catching whales don't know about it.
7	M	オーストラリアではくじらのにくを たべないのですか？ [Oosutoraria-dewa kujira-no niku-o tabenaino desu ka?]	Don't you eat whale in Australia?
8	L	問題は日本とノルウェーしかたべないよ [Mondai-wa nihon-go noruwee-shika tabenai yo]	The problem is that no one but Japanese and Norwegian eats [whale].
9	L	アメリカとかオーストラリアはほげいをやめてほし [Amerika toka Oosutoria-wa hogee-o yamete hoshi]	America and Australia want to stop whaling.
10	M	でも 日本でもすーぱーとかで くじらのにくを うってるの みたことないよ [Demo nihon-demo suupaa toka-de kujira-no niku-o utteruno mitakoto nai yo]	But I have never seen whale meat on sale at the supermarket in Japan.
11	L	けど日本人はくじらのにくをまだたべたい [Kedo nihonjin-wa kujira-no niku-o mada tabetai]	But Japanese still want to eat whale.
12	M	むかし せかいで ほげいがきんしされたよね [mukashi sekai-de hogee-ga kinshi saretai yo ne]	Long time ago, whaling was banned in the world.
13	L	その時オーストラリアもほげいしたよ [sono toki Oosutoraria-mo hogee shita yo]	At that time, Australia was also practicing whaling.
14	M	くじらにくの せんもんてんみたいのはすこしあるよ [kujira niku-no senmonten mitai no-wa sukoshi aru yo]	There are small numbers of shops specialized for [selling] whale.
15	L	けどにくたべてない けどにくはたべてない [kedo niku tabete nai kedo niku-wa tabete nai]	But we didn't eat meat, but didn't eat meat.
16	L	くじらのあぶらをつかいました [kujira-no abura-o tsukaimashita]	We used whale oil.
17	M	ほげいして どうしたの？ がいこくに うったの？ [hogee shite doo shita no? Gaikoku-ni utta no??]	So what did you do after catching whale? Did you sell it to foreign countries?
18	M	なんのために あぶらをつかったの？ [nan-no tame-ni abura-o tsukatta no?]	What did you use [whale] oil for?
19	L	あぶらのlamp [abura-no lamp]	Oil lamp.
20	M	いま ほげいしているくにが すくないから、くじらが ふえすぎているってきいたよ [ima hogee shite iru kuni-ga sukunai kara kujira-ga fuesugite irutte kiita yo]	I heard that the number of whales is increasing excessively because there aren't many countries which practice whaling.
21	M	ああ あぶらのランプね [aa abura-no ranpu ne]	Uh oil lamp.

22	L	そうよ [soo yo]	That's right.
23	M	なんで ほげいをやめてほしいの？ [nan-de hogee-o yamete hoshii no?]	Why do you want to stop whaling?
24	M	くじらがすきな の？ [Kujira-ga sukina no?]	Do you like whales?
25	L	だから私の意見は一年で日本じ n [dakara watashi no iken-wa ichinen-de nihonjin]	So my opinion is in one year Japanese
26	L	は 1 0 0 0 とか 2 0 0 0 くじらをつかまてもいい [wa 1000 toka 2000 kujira-o tsukamaete mo ii]	Is OK to catch 1000 or 2000 whales.
27	M	うん。いま 日本は それいじょうの くじらを つかまえてるの？ [un. Ima nihon-wa sore ijyoo-no kujira-o tsukamaeteru no??]	OK. Does Japan catch whales more than that now?

References

- Ågren, M. 2009. Morphological development in written L2 French: A Processability perspective. In *Research in Second Language Acquisition: Empirical Evidence Across Languages*, D. Keatinge & J.-U. Keßler (eds), 121-151. Newcastle, UK, Cambridge Scholars Publishing.
- Aissen, J. 2003. Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory* 213, 435-448. *Glossologia* 39 (2-3), 7-20.
- Alsina, A. 1996. *The Role of Argument Structure in Grammar*. Stanford, CA, CSLI Publications.
- Anđelković, D. 2000. Razumevanje imeničke morfologije srpskog jezika na predškolskom uzrastu [The acquisition of inflectional case system in Serbian]. *Psihologija* 33 (1-2), 145-170.
- Andersen, R.W. 1984. The one-to-one principle of interlanguage construction. *Language Learning* 34, 77-95.
- Andersen, R.W. 1993. Four operating principles and input distribution as explanations for underdeveloped and mature morphological systems. In *Progression and regression in language*, K. Hyltenstam & A. Viborg (eds), 309-339. New York, Cambridge University Press.
- Andorno, C., Bernini, G., Giacalone Ramat, A. & Valentini A. 2003. Sintassi della frase semplice. In *Verso l'Italiano. Percorsi e Strategie di Acquisizione*, A. Giacalone Ramat (ed.), 116-154. Roma, Carocci.
- Andrews, A. & Manning, C. 1999. *Complex Predicates and Information Spreading in LFG*. Stanford, CA, CSLI Publications.
- APA. 1994. *Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC, American Psychiatric Association.
- APA. 2013. Diagnostic and statistical manual of mental disorders, 5th edition, and clinical utility. *The Journal of nervous and mental disease* 2019 (2), 727. doi: 10.1097/NMD.0b013e3182a2168a.
- Artoni, D. & Magnani, M. 2013. LFG contribution in second language acquisition research: The development of case in Russian L2. In *Proceedings of the LFG13 Conference*. M. Butt & T.H. King (eds), 69-89. Stanford, CA, CSLI Publications.
- Asudeh, A. 2004. *Resumption as Resource Management*. PhD thesis, Stanford University, CA.
- Asudeh, A. & Toivonen, I. 2010. Lexical-Functional Grammar. In *The Oxford Handbook of Linguistic Analysis*, B. Heine & H. Narrog (eds), 425-458. Oxford & New York, Oxford University Press.

- Auer, P. (ed.) 1998. *Code-Switching in Conversation: Language, Interaction and Identity*. London, Routledge.
- Avrutin, S. & Rohrbacher, B. 1997. Null subjects in Russian inverted constructions. *Proceedings of the Fourth Annual Workshop on Formal Approaches to Slavic Linguistics*. Ann Arbor, MA, Slavic Publications.
- Backhouse, A.E. 1993. *The Japanese Language: An Introduction*. Oxford, Oxford University Press.
- Baddeley, A. 2007. *Working Memory, Thought, and Action*. Oxford, Oxford University Press.
- Bailey, N., Madden, C., & Krashen, S. 1974. Is there a “natural sequence” in adult second language learning? *Language Learning* 24 (2), 235-243.
- Baron-Cohen, S. 2009. Autism: The Empathizing-Systemizing (E-S) Theory. *Annals of the New York Academy of Sciences* 1156, 68-80.
- Bartolucci, G., Pierce, S. & Streiner, D. 1980. Cross-sectional studies of grammatical morphemes in autistic and mentally retarded children. *Journal of Autism and Developmental Disorders* 10 (1), 39-50.
- Baten, K. 2011. Processability Theory and German. Case acquisition. *Language Learning* 61 (2), 455-505.
- Baten, K. 2013. *The Acquisition of the German Case System by Foreign Language Learners*. Amsterdam, Benjamins.
- Bates, E. 1976. *Language and Context: The Acquisition of Pragmatics*. New York, Academic Press.
- Bates, E. & MacWhinney, B. 1989. Functionalism and the Competition Model. In *The Cross-linguistic study of sentence processing*, B. MacWhinney & E. Bates (eds), 3-73. New York, Cambridge University Press.
- Berman, R. & Slobin D.I. 1994. *Relating Events in Narrative; Vol 1, A Crosslinguistic Developmental Study*. Hillsdale, NJ, Erlbaum.
- Best, C.T., Tyler, M.D., Gooding, T.N., Orlando, C.B. & Quann, C.A. 2009. Development of phonological constancy: Toddlers’ perception of native- and Jamaican-accented words. *Psychological Science* 20 (5), 539-542.
- Bettoni, C. 1991. Language variety among Italians: Anglicisation, attrition, attitudes. In *Language in Australia*, S. Romaine (ed.), 263-269. Cambridge, Cambridge University Press.
- Bettoni C. & Di Biase B. 2005. Sviluppo obbligatorio e progresso morfosintattico, un caso di Processabilità in italiano L2. *ITALS. Didattica e Linguistica dell’Italiano come Lingua Straniera* 2 (1), 27-48.
- Bettoni, C. & Di Biase, B. 2011. Beyond canonical order. The acquisition of marked word orders in Italian as a second language. In *EUROSLA Yearbook 11*, L. Roberts, G. Pallotti & C. Bettoni (eds), 244-272. Amsterdam, Benjamins.
- Bettoni, C., Di Biase, B. & Ferraris, S. 2008. Sviluppo sintattico e sviluppo morfologico: ipotesi di corrispondenze nella Processabilità dell’italiano L2. In *Competenze Lessicali e Discorsive nell’Acquisizione di Lingue Seconda*, G. Bernini, L. Spreafico & A. Valentini (eds), 355-382. Perugia, Guerra.

- Bettoni, C., Di Biase, B. & Nuzzo, E. 2009. Postverbal subject in Italian L2 – a Processability Theory approach. In *Research in Second Language Acquisition: Empirical Evidence Across Languages*, D. Keatinge & J.-U. Kessler (eds), 153-173. Newcastle, UK, Cambridge Scholars Publishing.
- Bettoni, C. & Fratter, I. 2013. Acquiring reflexive constructions in Italian as a second language – A Processability Theory account. Paper presented at the 23rd Annual Conference, European Second Language Association (EUROSLA 23), Amsterdam, 28-31 September 2013.
- Bever, T.G. 1970. The cognitive basis for linguistic structures. In *Cognition and the Development of Language*, J.R. Hayes (ed.), 279-362. New York: Wiley.
- Bialystok, E. & Craik, F.I.M. 2010. Cognitive and linguistic processing in the bilingual mind. *Current Directions in Psychological Science* 19 (1), 19-23.
- Blake, B. J. 1994. *Case*. Cambridge, Cambridge University Press.
- Blake, R. 2000. Computer mediated communication: A window on L2 Spanish interlanguage. *Language Learning and Technology* 4 (1), 120-136.
- Bock, K. & Levelt, W. 1994. Language production. Grammatical encoding. In *Handbook of Psycholinguistics*, M.A. Gernsbacher (ed.), 945-984. New York, Academic Press.
- Bonilla, C.L. 2012. Testing Processability Theory in L2 Spanish: Can readiness or markedness predict development? Doctoral dissertation, University of Pittsburgh, PA.
- Bonilla, C.L. 2014. From number agreement to the subjunctive: Evidence for Processability Theory in L2 Spanish. *Second Language Research*, 0267658314537291.
- Bonilla, C.L. 2015. From number agreement to the subjunctive: Evidence for Processability Theory in L2 Spanish. *Second Language Research* 31 (1), 53-74.
- Boss, B. 1996. German grammar for beginners: The Teachability Hypothesis and its relevance to the classroom. In *Who's Afraid of Teaching Grammar? (The University of Queensland Papers in Language and Linguistics, vol. 1)*, C. Arbon'es-Sol'a, J. Rolin-Iantzi & R. Sussex (eds), 93-100. Brisbane, University of Queensland Press.
- Boss, B. 2004. Wann ich habe Zeit, ich koche gern: Zum Erwerb der Inversion und Nebensatzwortstellung durch australische Deutschstudierende. *Deutsch als Fremdsprache* 41, 28-32.
- Bossong, G. 1983-1984. Animacy and Markedness in Universal Grammar. *Glossologia* 2-3, 7-20.
- Bossong, G. 1991. Differential object marking in Romance and beyond. In *New Analyses in Romance Linguistics. Selected Papers from the XVIII Linguistic Symposium on Romance Languages*, D. Wanner & D. Kibbee (eds), 143-170. Amsterdam, Benjamins.
- Boucher, J. 2012. Research review: Structural language in autistic spectrum disorder – characteristics and causes. *Journal of Child Psychology and Psychiatry* 53 (3), 219-233. doi: 10.1111/j.1469-7610.2011.02508.x.
- Bower, J. & Kawaguchi, S. 2011. Negotiation of meaning and corrective feedback in Japanese/English eTandem article. *Language Learning and Technology* 15 (1), 41-71.
- Bowles, M. & Montrul, S. 2009. Instructed L2 acquisition of Differential Object Marking in Spanish. In *Little Words: Their History, Phonology, Syntax, Semantics, Pragmatics, and*

- Acquisition*, R.P. Leow, H. Campos & D. Lardiere (eds), 199-209. Washington, Georgetown University Press.
- Bresnan, J. 1998. Pidgin genesis in Optimality Theory. In *Proceedings of the LFG98 Conference, Brisbane*, M. Butt & T. Holloway King (eds), 1-18. Stanford, CA, CSLI Publication, www.cslipublications.stanford.edu/LFG/3/lfg98.html.
- Bresnan, J. 2000. Optimal Syntax. In *Optimality Theory: Phonology, Syntax, and Acquisition*, J. Dekkers, F. van der Leew & J. van de Weijer (eds), 334-385. Oxford, Oxford University Press.
- Bresnan, J. 2001. *Lexical-Functional Syntax*. Oxford, Blackwell.
- Bresnan, J. & Mchombo, S. 1987. Topic, pronoun and agreement in Chichewa. *Language* 63, 741-782.
- Brigance, A.H. 1991. *Brigance K & 1 Screen for Kindergarten and First Grade – Revised*. North Billerica, MA, Curriculum Associates.
- Brown, R. 1973. *A First Language*. Harvard, MA, Harvard University Press.
- Browne, W. 1993. Serbo-Croat. In *The Slavonic Languages*, B. Comrie & G.G. Corbett (eds), 306-387. London, Routledge.
- Carpenter, P.A. & Just, M.A. 1989. The role of working memory in language comprehension. In *Complex Information Processing*, D. Klahr & K. Krovsky (eds), 31-68. Hillsdale, NJ, Erlbaum.
- Caselli, C.M., Leonard, L.B., Volterra, V. & Campagnoli, M.G. 1993. Toward mastery of Italian morphology: A cross-sectional study. *Journal of Child Language* 20, 377-393.
- CEFR 2001. *Modern Languages: Learning, Teaching, Assessment. A Common European Framework of Reference*. Cambridge, Cambridge University Press.
- Chambers, A. & Bax, S. 2006. Making CALL work: Towards normalization. *System* 34, 465-479.
- Charters, H., Dao, L. & Jansen, L. 2011. Reassessing the applicability of Processability Theory: The case of nominal plural. *Second Language Research* 27 (4), 509-533.
- Cheng, Lai-Shen L. & Corver, N. (eds) 2006. *Wh-Movement: Moving On*. Cambridge, MA, MIT Press.
- Chini, M. & Ferraris, S. 2003. Morfologia del nome. In *Verso l'Italiano. Percorsi e Strategie di Acquisizione*, A. Giacalone Ramat (ed.), 37-69. Roma, Carocci.
- Choi, H.-W. 1999. *Optimizing Structure in Context: Scrambling and Information structure*. Stanford, CA, CSLI Publications.
- Choi, H.-W. 2001. Phrase structure, information structure, and resolution of mismatch. In *Formal and Empirical Issues in Optimality Theoretic Syntax*. P. Sells (ed.), 17-62. Stanford, CSLI Publications.
- Chomsky, N. 1977. On Wh-movement. In *Formal Syntax*, P.W. Culicover, T. Wasow & A. Akmajian (eds), 71-133. New York, Academic Press.
- Clahsen, H. 1980. Psycholinguistic aspects of L2 acquisition. In *Second Language Development*, S. Felix, (ed.), 57-79. Tübingen, Narr.
- Clahsen, H. 1984. The acquisition of German word order: A test case for cognitive approaches to L2 development. In *Second Language*, R.W. Andersen (ed.), 219-242. Rowley, Newbury House.

- Clahsen, H., Meisel, J. & Pienemann, M. 1983. *Deutsch als Zweitsprache. Der Spracherwerb ausländischer Arbeiter*. Tübingen, Narr.
- Clahsen, H. & Felser, C. 2006. Grammatical processing in languagelearners. *Applied Psycholinguistics* 27, 3-42.
- Clyne, M. 2003. *Dynamics of Language Contact*. Cambridge, Cambridge University Press.
- Comrie, B. 1979. Definite and animate direct objects: A natural class. *Linguistica Silesiana* 3, 13-21.
- Comrie, B. 1989. *Language Universals and Linguistic Typology*. Oxford, Blackwell.
- Comrie, B. 2011. Russian. In *The Major Languages of the World*, B. Comrie (ed.), 274-288. London & New York, Routledge.
- Conroy, L. 2007. Extended Processability Theory and its application to L2-Swedish. In *Proceedings, In Between Wor(l)ds: Transformation and Translation*, University of Melbourne. <http://repository.unimelb.edu.au/10187/2213>
- Cook, V. 1993. *Linguistics and Second Language Acquisition*. New York, St. Martin's.
- Coonrod, E.E. & Stone, W.L. 2005. Screening for autism in young children. In *Handbook of Autism and Pervasive Developmental Disorders*, F.R. Volkmar, R. Paul, A. Klin & D. Cohen (eds), 3rd ed., vol. 1, 707-729. Hoboken, NJ, Wiley & Sons.
- Corbett, G. & Browne, W. (2011). Serbo-Croat: Bosnian, Croatian, Montenegrin, Serbian. In *The Major Languages of the World*, B. Comrie (ed.), 330-346. London & New York, Routledge.
- Cziko, G. 2004. Electronic tandem language learning (eTandem): A third approach to second language learning for the 21st century. *CALICO Journal* 22 (1), 25-39.
- Dalrymple, M. 2001. *Syntax and Semantics: Lexical Functional Grammar*. San Diego, CA, Academic Press.
- Dalrymple, M. & Nikolaeva, I. 2011. *Objects and Information Structure*. Cambridge, Cambridge University Press.
- Dardano, M. & Trifone, P. 1997. *La Nuova Grammatica della Lingua Italiana*. Bologna, Zanichelli.
- De Bot, K. 1992. A bilingual production model: Levelt's 'Speaking' model adapted. *Applied Linguistics* 13 (1), 1-24.
- De Houwer, A. 1990. *The Acquisition of Two Languages from Birth: A Case Study*. Cambridge, Cambridge University Press.
- De Houwer, A. 2005. Early bilingual acquisition: Focus on morphosyntax and the Separate Development Hypothesis. In *Handbook of Bilingualism: Psycholinguistic Approaches*, J.F. Kroll & A.M.B. De Groot (eds), 30-48. Oxford, Oxford University Press.
- DeKeyser, R.M. 2005. What makes learning second-language grammar difficult? A review of issues. *Language Learning* 55, 1-25.
- Di Biase, B. 1998. Processability Theory and the acquisition of Italian L2. Paper delivered at the *European Science Foundation Conference: The Structure of Learner Language, Acquafredda di Maratea, Italy, September 26-October 1, 1998*.
- Di Biase, B. 2000. Preface. In *Stages of Development for English as a Second Language*, M. Johnston, v-vii. Sydney, LARC, University of Western Sydney.

- Di Biase, B. 2002. Focusing strategies in second language development: A classroom-based study of Italian L2 in primary school. In *Developing a Second Language: Acquisition, Processing and Pedagogy of Arabic, Chinese, English, Italian, Japanese, Swedish*, B. Di Biase (ed.), 95-120. Melbourne, Language Australia.
- Di Biase, B. 2005. The Topic Hypothesis in Processability Theory. Paper presented at the *5th International Symposium on Processability, Second Language Acquisition and Bilingualism*. Deakin University, Melbourne, 26-28 September, 2005.
- Di Biase, B. 2007. *A Processability Approach to the Acquisition of Italian as a Second Language: Theory and Applications*. PhD thesis, Australian National University, Canberra, <http://hdl.handle.net/1885/6982>.
- Di Biase, B. 2008. Focus-on-form and development in L2 learning. In *Processing Approaches to Second Language Development and Second Language learning*, J.-U. Kefler (ed.), 197-220. Newcastle, UK, Cambridge Scholars Publishing.
- Di Biase, B. & Bettoni, C. 2007. Funzioni discorsive e Processabilità in italiano. In *Imparare una Lingua. Recenti Sviluppi Teorici e Proposte Applicative*, M. Chini, A. De Meo, P. Desideri & G. Pallotti (eds), 209-233. Perugia, Guerra.
- Di Biase, B. & Bettoni, C. 2015. Subject realisation in Italian L2, a cross-sectional study of production data. In *Sintassi. Atti del 46° Congresso Internazionale di Studi della Società di Linguistica Italiana, Siena, 27-29.09.2012*, C. Bruno, S. Casini, F. Gallina & R. Siebetchu (eds), 589-604. Roma, Bulzoni.
- Di Biase, B. & Hinger, B. 2011. Topic Hypothesis in Processability Theory: The case of Spanish. Paper presented at *The 11th Symposium on Processability Approaches to Language Acquisition (PALA)*. Innsbruck. <http://www.slashdocs.com/mwppqry/topic-hypothesis-in-processability-theory-di-biase-hinger-2011.html>
- Di Biase, B. & Kawaguchi, S. 2002. Exploring the typological plausibility of Processability Theory: Language development in Italian second language and Japanese second language. *Second Language Research* 18 (3), 272-300.
- Di Biase, B. & Kawaguchi, S. 2012. Processability Theory. In *Encyclopedia of Second Language Acquisition*, P. Robinson (ed.), 512-517. London, Routledge.
- Diessel, H. 2004. *The Acquisition of Complex Sentences*. Cambridge: Cambridge University Press.
- Dik, S.C. 1997. *The Theory of Functional Grammar*. Berlin & New York, Mouton de Gruyter.
- Dimitrijević, J. 2004a. Language shift and contact-induced change: Serbian in contact with English in Australia. *Melbourne Papers in Linguistics and Applied Linguistics* 2, 1-19.
- Dimitrijević, J. 2004b. Code-switching: Structure and meaning. *Facta Universitatis* 3 (1), 37-46.
- Dimitrijević-Savić, J. 2008. Convergence and attrition: Serbian in contact with English in Australia. *Journal of Slavic Linguistics* 16 (1), 57-90.
- Doi, T. & Yoshioka, K. 1987. Which grammatical structures should be taught when? Implications on the Pienemann-Johnston model for teaching Japanese as a foreign language. Paper presented at the *9th Hawaii Association of Teachers of Japanese*. University of Hawaii at Manoa, Honolulu.

- Doi, T. & Yoshioka, K. 1990. Jyoshi no shuutoku ni okeru gengo unyoojoo no seiyaku –Pienemann-Johnston moderu no nihongo shuutoku kenkyuu e no ooyoo. *Proceedings of 1st Conference on SLA and Teaching* 1, 23-33.
- Donaldson, R.P. & Kotter, M. 1999. Language learning in cyberspace: Teleporting the classroom into the target culture. *CALICO Journal* 16 (4), 531-557.
- Doughty, C.J. 1991. Second language instruction does make a difference: Evidence from an empirical study of ESL relativization. *Studies in Second Language Acquisition* 13 (4), 431-469.
- Doughty, C.J. & Long, M.H. (eds) 2003. *The Handbook of Second Language Acquisition*. Oxford, Blackwell.
- Dryer, M.S. 1989. Large linguistic areas and language sampling. *Studies in Language* 13, 257-92.
- Dryer, M.S. 2013. Determining Dominant Word Order. In *The World Atlas of Language Structures Online*, M.S. Dryer & M. Haspelmath (eds), ch. 26. Leipzig, Max Planck Institute for Evolutionary Anthropology. Available online at <<http://wals.info/chapter/26>>, accessed on 19-03-2015.
- Dulay, H. & Burt, M. 1974. Natural sequences in child second language acquisition. *Language Learning* 24 (1), 37-53.
- Dyson, B. 2009. Processability Theory and the role of morphology in English as a second language development: a longitudinal study. *Second Language Research* 25 (3), 355-376.
- Eigsti, I.-M., Bennetto, L. & Dadlani, M. 2007. Beyond pragmatics: Morphosyntactic development in autism. *Journal of Autism and Developmental Disorders* 37 (6), 1007-1023.
- Ellis, N. & Collins, L. 2009. Input and second language acquisition: The roles of frequency, form, and function. *The Modern Language Journal* 93, 329-335.
- Ellis, R. 1989. Are classroom and naturalistic acquisition the same? A study of the classroom acquisition of German word order rules. *Studies in Second Language Acquisition* 11, 305-328.
- Ellis, R. 1994. *The Study of Second Language Acquisition*. Oxford, Oxford University Press.
- Ellis, R. 2008. *The Study of Second Language Acquisition*. Second Edition. Oxford, Oxford University Press.
- Erteschik-Shir, N. 2007. *Information Structure. The Syntax-Discourse Interface*. New York, Oxford University Press.
- Falk, Y.N. 2001. *Lexical-Functional Grammar: An Introduction to Parallel Constraint-based Syntax*. Stanford, CA, CSLI Publication.
- Falk, Y.N. 2006. *Subjects and Universal Grammar: An Explanatory Theory*. Cambridge, Cambridge University Press.
- Farley, A.P. & McCollam, K. 2004. Learner readiness and L2 production in Spanish: Processability Theory on trial. *Estudios de Lingüística Aplicada* 40, 47-69.
- Fava, E. 1995. Tipi di frasi principali. Il tipo interrogativo. In *Grande Grammatica Italiana di Consultazione*, L. Renzi, G. Salvi & A. Cardinaletti (eds), vol. 3, 70-127. Bologna, Il Mulino.

- Filipek, P.A., Accardo, P.J., Baranek, G.T., Cook, E.H., Dawson, G. & Gordon, B. 1999. The screening and diagnosis of autistic spectrum disorders. *Journal of Autism and Developmental Disorders* 29 (6), 439-484.
- Finegan, E. 2011. English. In *The Major Languages of the World*, B. Comrie (ed.), 59-85. London & New York, Routledge.
- Franck, J., Vigliocco, G. & Nicol, J. 2002. Subject-verb agreement errors in French and English: The role of syntactic hierarchy. *Language and Cognitive Processes* 17, 371-404.
- Franks, S. 1994. Parametric properties of numeral phrases in Slavic. *Natural Language and Linguistic Theory* 12, 597-674.
- Frey, W. 2004. A medial topic position for German. *Linguistische Berichte* 198, 153-190.
- Gass, S.M. 1994. The reliability of second-language grammaticality judgments. In *Research Methodology in Second Language Acquisition*, E.E. Tarone, S.M. Gass & A.D. Cohen (eds), 302-322. Hillsdale, NJ, Erlbaum.
- Gass, S.M. & Mackey, A. 2012. *The Routledge Handbook of Second Language Acquisition*. New York, Routledge.
- Gass, S.M. & Selinker, L. 2001. *Second Language Acquisition. An Introductory Course*. 2nd edition. Hillsdale, NJ, Erlbaum.
- Givón, T. 1983. Topic continuity and word-order pragmatics in Ute. In *Topic Continuity in Discourse*, T. Givón (ed.), 145-195. Amsterdam, Benjamins.
- Givón, T. 2001. *Syntax*. Vols 1-2. Amsterdam, Benjamins.
- Glahn, E., Håkansson, G., Hammarberg, B., Holmen, A., Hvenekilde, A. & Lund, K. 2001. Processability in Scandinavian second language acquisition. *Studies in Second Language Acquisition* 25, 389-416.
- Green, J.N. 2011. Spanish. In *The World's Major Languages*, B. Comrie (ed.), 2nd ed., 197-216. London & New York, Routledge.
- Greenberg, J.H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of Language*, J.H. Greenberg (ed.), 58-90. Cambridge, MA, MIT Press.
- Grosjean, F. & Li P. 2013. *The Psycholinguistics of Bilingualism*. Malden, MA, Wiley-Blackwell.
- Guijarro-Fuentes, P. 2011. Feature composition in Differential Object Marking. *EUROSLA Yearbook* 11, L. Roberts, G. Pallotti & C. Bettoni (eds), 138-164. Amsterdam, Benjamins.
- Guijarro-Fuentes, P. 2012. The acquisition of interpretable features in L2 Spanish: Personal *a*. *Bilingualism: Language and Cognition* 15, 701-720.
- Habermehl, S. 2005. *Der Erwerb der Verbstellungsregeln in der Zweitsprache Deutsch durch Kinder mit Russischer und Türkischer Muttersprache*. Tübingen, Max Niemeyer.
- Håkansson, G. 2001. Tense morphology and Verb-second in Swedish L1 children, L2 children and children with SLI. *Bilingualism: Language and Cognition* 4, 85-99.
- Håkansson, G. 2005. Similarities and differences in L1 and L2 development: Opening up the perspective: Including SLI. In *Cross-linguistic Aspects of Processability Theory*, M. Pienemann (ed.), 179-198. Amsterdam, Benjamins.
- Håkansson, G. & Norrby, C. 2007. Processability Theory applied to written and oral

- Swedish. In *Second Language Acquisition Research: Theory-construction and Testing*, F. Mansouri (ed.), 81-94. Newcastle, UK, Cambridge Scholar Publishing.
- Håkansson, G., Pienemann, M. & Sayehli, S. 2002. Transfer and typological proximity in the context of second language processing. *Second Language Acquisition Research* 18 (3), 250-273.
- Håkansson, G., Salameh, E. & Nettelbladt, U. 2003. Measuring language development in bilingual children: Swedish-Arabic children with and without language impairment. *Linguistics* 41, 255-288.
- Hambly, C., & Fombonne, E. 2012. The impact of bilingual environments on language development in children with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders* 42 (7), 1342-1352. doi: 10.1007/s10803-011-1365-z.
- Hambly, C., & Fombonne, E. 2014. Factors influencing bilingual expressive vocabulary size in children with autism spectrum disorders. *Research in Autism Spectrum Disorders* 8(9), 1079-1089. doi: <http://dx.doi.org/10.1016/j.rasd.2014.05.013>.
- Hammond, L. 2005. *Serbian: An Essential Grammar*. London, Routledge.
- Happé, F. & Frith, U. 2006. The weak coherence account, Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders* 36 (1), 5-25.
- Harada, S. 1971. Ga-no conversion and idiolectal variations in Japanese. *Gengo Kenkyuu* 60, 25-38.
- Hatch, E. & Lazaraton, A. 1991. *The research manual: Design and Statistics for Applied Linguistics*. Boston, MA, Heinle & Heinle.
- Hayashi, A., Tomlin, R.S. & Yokota, T. 2002. Attention detection and Japanese. Unpublished manuscript.
- Hinger, B. 2011. Sprache lehren – Sprache überprüfen – Sprache erwerben: Empirische Einsichten in den schulischen Spanischunterricht – eine Fallstudie. Habilitation thesis, University of Innsbruck, Innsbruck.
- Hopp, H. 2007. Cross-linguistic differences at the syntax-discourse interface in off- and on-line L2 performance. In *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA)*, A. Belikova & al. (eds), 147-158. Somerville, MA, Cascadia Proceedings Project.
- Hopper, J.P. & Thompson, S.A. 1980. Transitivity in grammar and discourse. *Language* 56, 251-299.
- Horvath, J. 1986. *FOCUS in the Theory of Grammar and Syntax in Hungarian*. Dordrecht, Foris.
- Howlin, P. 1984. The acquisition of grammatical morphemes in autistic children: A critique and replication of the findings of Bartolucci, Pierce, and Streiner, 1980. *Journal of Autism and Developmental Disorders* 14 (2), 127-136.
- Howlin, P., Goode, S., Hutton, J. & Rutter, M. 2004. Adult outcome for children with autism. *Journal of Child Psychology and Psychiatry* 45 (2), 212-229.
- Hualde, J.I., Olarra, A. & Escobar, A.M. 2001. *Introducción a la Lingüística Hispánica*. Cambridge, Cambridge University Press.
- Hudry, K., Chandler, S., Bedford, R., Pasco, G., Gliga, T., Elsabbagh, M., ... Charman, T. 2014. Early language profiles in infants at high-risk for Autism Spectrum Disorders.

- Journal of Autism and Developmental Disorders* 44(1), 154-167. doi: 10.1007/s10803-013-1861-4.
- Huter, K. 1996. *Atarashii no Kuruma* and other old friends – the acquisition of Japanese syntax. *Australian Review of Applied Linguistics* 19, 39-60.
- Huter, K. 1997. “*Onnanohito wa duressu desu. That’s why the lady is a dress*”, Developmental stages in Japanese second language acquisition. *Australian Studies in Language Acquisition* 6, 1-39.
- Ishikawa, A. 1985. *Complex Predicates and Lexical Operations in Japanese*. Doctoral dissertation, Stanford University, CA.
- Itani-Adams, Y. 2007. Lexical and grammatical development in Japanese-English bilingual first language acquisition. In *Second Language Acquisition Research: Theory Construction and Testing*, Mansouri, F. (ed.), 173-198. Newcastle, UK., Cambridge Scholars.
- Itani-Adams, Y. 2008. *One Child, Two Languages: Bilingual First Language Acquisition in Japanese and English*. PhD thesis, University of Western Sydney, Sydney.
- Itani-Adams, Y. 2009. Development of discourse functions in Japanese and English bilingual first language acquisition. In *Research in Second Language Acquisition: Empirical Evidence Across Languages*, J.-U. Kessler & D. Keatinge (eds), 41-66. Newcastle, UK, Cambridge Scholars Publishing.
- Itani-Adams, Y. 2011. Bilingual first language acquisition. In *Studying Processability Theory*, M. Pienemann & J.U. Kessler (eds), 121-130. Amsterdam, Benjamins.
- Itani-Adams, Y. 2013. *One Child and Two Languages: Acquisition of Japanese and English as Bilingual First Languages*. München, Lincom.
- Iwasaki, J. 2008. Acquiring Japanese as a second language (JSL) in a naturalistic context: A longitudinal study of a young child from a Processability Theory (PT) perspective. In *Second Language Acquisition and the Younger Learner: Child’s Play?* J. Philp, R. Oliver & A. Mackey (eds), 231-253. Amsterdam, Benjamins.
- Iwasaki, J. 2013. A Processability Theory (PT)-based analysis of the acquisition of Japanese morphology and syntax: A case of an intensive adult learner. *Second Language* 12, 21-42.
- Iwasaki, J. & Oliver, R. 2003. Chat-line interaction and negative feedback. *Australian Review of Applied Linguistics* 17, 60-73.
- Jackendoff, R.S. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, MA, MIT Press.
- Jakobson, R. 1959. Boas’ view of grammatical meaning. *American Anthropologist* 61, 139-145.
- Jakobson, R. 1971. Beitrag zur allgemeinen Kasuslehre. In *Selected Writings II*, R. Jakobson (ed.), 23-71. The Hague, Mouton.
- Jansen, L. 2008. Acquisition of German word order in tutored learners: A cross-sectional study in a wider theoretical context. *Language Learning* 58, 185-231.
- Jarrold, C., Boucher, J. & Russell, J. 1997. Language profiles in children with autism: Theoretical and methodological implications. *Autism* 1 (1), 57-76.
- Jellinek, M.S., Toppelberg, C.O., Snow, C.E., & Teger-Flusberg, H. 1999. Severe developmental disorders and bilingualism. *Journal of the American Academy of Child*

- J. Adolescent Psychiatry* 38(9), 1197-1199. doi: <http://dx.doi.org/10.1097/00004583-199909000-00027>
- Jia, G. & Fuse, A. 2007. Acquisition of English grammatical morphology by native Mandarin-speaking children and adolescents: Age-related differences. *Journal of Speech, Language and Hearing Research* 50, 1280-1299.
- Johnston, M. 1985a. *Syntactic and morphological progressions in learner English*. Canberra, Department of Immigration and Ethnic Affairs.
- Johnston, M. 1985b. SLA Research in the Adult Migrant Education Program. *Prospect: A Journal of Adult Migrant Education* 1 (1), 19-46.
- Johnston, M. 1987. Understanding learner language. In *Applying Second Language Acquisition Research*, D. Nunan (ed.), 2-44. Adelaide, National Curriculum Research Centre, Adult Migrant Education Program.
- Johnston, M. 1994. *Second Language Acquisition: A Classroom Perspective*. Sydney, NLLIA Language Acquisition Research Centre, University of Western Sydney.
- Johnston, M. 1995. Stages of Acquisition of Spanish as a Second Language. *Australian Studies in Language Acquisition* 4, 1-28.
- Johnston, M. 1997. Development and Variation in Learner Language. PhD thesis, Australian National University, Canberra.
- Johnston, M. 2000. *Stages of development for English as a Second Language*. Melbourne, Language Australia.
- Jordan, G. 2004. *Theory Construction in Second Language Acquisition*. Amsterdam, Benjamins.
- Kageyama, T. 1999. Word formation. In *The Handbook of Japanese Linguistics*, N. Tsujimura (ed.), 297-325. Malden, MA, Blackwell.
- Kahnemann, D. 1973. *Attention and Effort*. Englewood Cliffs, NJ, Prentice-Hall.
- Kalvestinova, E.D. 2007. Aspects of word order in Russian. PhD dissertation, University of Iowa, IA.
- Kanner, L. 1943. Autistic disturbances of affective contact. *Nervous Child* 2 (2), 217-250.
- Kanner, L. 1971. Follow-up study of eleven autistic children originally reported in 1943. *Journal of Autism and Developmental Disorders* 1 (2), 119-145.
- Kaplan, R.M. 1995. The formal architecture of lexical-functional grammar. In *Formal Issues in Lexical-Functional Grammar*, M. Dalrymple, R.M. Kaplan, J.T. Maxwell III & A. Zaenen (eds), 7-27. Stanford University, CSLI.
- Kaplan, R.M. & Bresnan, J. 1982. Lexical-Functional Grammar: A formal system for grammatical representation. In *The Mental Representation of Grammatical Relations*, J. Bresnan (ed.), 173-281. Cambridge, MA, MIT Press.
- Kay-Raining Bird, E., Lamond, E., & Holden, J. 2012. Survey of bilingualism in autism spectrum disorders. *International Journal of Language & Communication Disorders* 47(1), 52-64. doi: 10.1111/j.1460-6984.2011.00071.x.
- Kawaguchi, S. 2000. Acquisition of Japanese verbal morphology: Applying Processability Theory to Japanese. *Studia linguistica* 54 (2), 238-248.
- Kawaguchi, S. 2002. Grammatical development in learners of Japanese as a second language. In *Developing a second language: Acquisition, processing and pedagogy of Arabic*,

- Chinese, English, Italian, Japanese, Swedish*, B. Di Biase (ed.), 17-28. Melbourne, Language Australia.
- Kawaguchi, S. 2005. Argument structure and syntactic development in Japanese as a second language. In *Cross-Linguistic Aspects of Processability Theory*, M. Pienemann (ed.), 253-299. Amsterdam, Benjamins.
- Kawaguchi, S. 2007. Lexical Mapping Theory and Processability Theory: A case study in Japanese. In *Second Language Acquisition Research: Theory-Construction and Testing*, F. Mansouri (ed.), 39-90. Newcastle, UK, Cambridge Scholars Publishing.
- Kawaguchi, S. 2008. Language typology and Processability Theory. In *Processing Approaches to Second Language Development and Second Language Learning*, J.-U. Kefler (ed.), 89-112. Newcastle, UK, Cambridge Scholars Publishing.
- Kawaguchi, S. 2009a. Acquiring causative constructions in Japanese as a second language, *The Journal of Japanese Studies* 29 (2), 273-291.
- Kawaguchi, S. 2009b. Acquisition of non-canonical order in Japanese as a second language: The case of causative structure. In *Research in Second Language Acquisition: Empirical Evidence Across Languages*, D. Keatinge & J.-U. Kefler (eds), 213-239. Newcastle, UK, Cambridge Scholars Publishing.
- Kawaguchi, S. 2010. *Learning Japanese as a Second Language. A Processability Perspective*. Amherst, NY, Cambria Press.
- Kawaguchi, S. 2011. Japanese as a second language: A test case for the typological plausibility of PT. In *Studying Processability Theory. An Introductory Textbook*, M. Pienemann & J.-U. Kefler (eds), 99-105. Amsterdam, Benjamins.
- Kawaguchi, S. 2013. The relationship between lexical and syntactic development in English as a second language. In *Language Acquisition and Use in Multilingual Contexts: Theory and Practice*, A.F. Mattisson & C. Norrby (eds), 92-106. Lund, Sweden, University of Lund, Travaux de L'Institut de Linguistique De Lund.
- Kawaguchi, S. (in press). Question constructions, argument mapping, and vocabulary development in English L2 by Japanese speakers: A cross-sectional study. In *Developing and Assessing Second Language Grammars across Languages*. J.-U. Kefler, A. Lenzig & M. Liebner (eds). Amsterdam, Benjamins.
- Kawaguchi, S. & Di Biase, B. 2005. Second language development at the syntax-pragmatics interface. Paper presented at the *Seventh Annual International Conference of the Japanese Society of Language Sciences, JSLS. Sophia University, Tokyo, June 25-26, 2005*.
- Kawaguchi, S., & Di Biase, B. 2012. Acquiring procedural skills in L2: Processability Theory and skill acquisition. *Studies in Language Sciences* 11, 70-99.
- Keatinge, D. 2008. Key issues in designing tasks for data elicitation purposes. Paper presented at *The 8th Symposium on Processability Approaches to Language Acquisition (PALA)*. Verona, 15-16 September 2008.
- Keatinge, D. & Kefler J.-U. 2009. The acquisition of the passive voice in L2 English: Perception and production. In *Research in Second Language Acquisition: Empirical Evidence Across Languages*, D. Keatinge & J.-U. Kefler (eds), 67-92. Newcastle, UK, Cambridge Scholars Publishing.
- Keenan, E.L. 1979. On surface form and logical form. *Studies in Linguistic Sciences* 8, 1-41.

- Keenan, E.O. & Comrie, B. 1977. Noun phrase accessibility and Universal Grammar. *Linguistic Inquiry* 8, 63-99.
- Kempe, V. & MacWhinney, B. 1998. The acquisition of case-marking by adult learners of Russian and German. Paper 218, Department of Psychology. Pittsburgh, PA, Carnegie Mellon University.
- Kempen, G. 1998. Comparing and explaining the trajectories of first and second language acquisition: In search of the right mix of psychological and linguistic factors. *Bilingualism: Language and Cognition* 1, 29-30.
- Kempen, G. & Hoenkamp, E. 1987. An incremental procedural grammar for sentence formulation. *Cognitive Science* 1 (1), 201-258.
- Keßler, J.-U. & Pienemann, M. 2011. Research methodology. In *Studying Processability Theory. An Introductory Textbook*. M. Pienemann & J.-U. Keßler (eds), 84-96. Amsterdam, Benjamins.
- Kim, J.-B. & Sells, P. 2008. *English Syntax: An Introduction*. Stanford, CA, CSLI Publications.
- Kim, S.H., Paul, R., Tager-Flusberg, H., & Lord, C. 2014. Language and Communication in Autism. In *Handbook of Autism and Pervasive Developmental Disorders*, F.R. Volkmar, S.J. Rogers, P. Rhea & K.A. Pelphrey (eds), 4th ed., vol. 1, 230-262. Hoboken, NJ, Wiley & Sons.
- King, T.H. 1995. *Configuring Topic and Focus in Russian*. Stanford, CA, CSLI Publications.
- Kjelgaard, M.M. & Tager-Flusberg, H. 2001. An investigation of language impairment in autism: Implications for genetic subgroups. *Language and Cognitive Processes* 16(2), 287-308.
- Klein, W. & Carroll, M. 1992. The acquisition of German. In *Utterance Structure. Developing Grammars Again*, W. Klein & C. Perdue (eds), 123-188. Amsterdam, Benjamins.
- Kormos, J. 2006. *Speech Production and Second Language Acquisition*. Mahwah, NJ, Erlbaum.
- Kotter, M. 2003. Negotiation of meaning and codeswitching in online tandems. *Language Learning and Technology* 7 (2), 145-172.
- Krashen, S. 1976. Formal and informal linguistic environments in language learning and language acquisition. *TESOL Quarterly* 10 (2), 157-168.
- Krashen, S. 1977. Some issues relating to the monitor model. In *Teaching and Learning English as a Second Language: Trends in Research and Practice: Selected Papers from the Eleventh Annual Convention of Teachers of English to Speakers of Other Languages, Miami, Florida, April 26 - May 1, 1977*, 144-158. Washington, DC, Teachers of English to Speakers of Other Languages.
- Krashen, S. 1979. A Response to McLaughlin, "The Monitor Model: Some Methodological Considerations". *Language Learning* 29, 151-67.
- Krashen, S. 1982. *Principles and Practice in Second Language Acquisition*. Oxford, Oxford University Press.
- Krashen, S. 1985. *The Input Hypothesis: Issues and Implications*. New York, Longman.
- Kroeger, P.R. 1993. *Phrase Structure and Grammatical Relations in Tagalog*. Stanford, CA, CSLI Publication.

- Kroeger, P.R. 2004. *Analysing Syntax: A Lexical-functional Approach*. Cambridge, Cambridge University Press.
- Kroeger, P.R. 2005. *Analysing Grammar: An Introduction*. Cambridge, Cambridge University Press.
- Kroll, J.F. & de Groot, A.M.B. (eds) 2005. *Handbook of Bilingualism: Psychological Approaches*. Oxford and New York, Oxford University Press.
- Kuno, S. 1973. *The Structure of the Japanese Language*. Cambridge, MA, MIT Press.
- Lai, C. & Zhao, Y. 2006. Noticing and text-based chat. *Language Learning and Technology* 10 (3), 102-120.
- Laca, B. 2006. El objeto directo. La marcación preposicional. In *Sintaxis historica de la lengua espanola*, C. Company Company (ed.), vol. 1, 421-75. Distrito Federal, Universidad Nacional Autónoma de México.
- Lambrecht, K. 1988. There was a farmer had a dog: Syntactic amalgams revisited. *Proceedings of the 14th Annual Meeting of the Berkeley Linguistics Society* 14, 319-339.
- Lambrecht, K. 1994. *Information Structure and Sentence Form: Topic, Focus and Mental Representation of Discourse Referents*. Cambridge, Cambridge University Press.
- Lardiere, D. 1998. Case and tense in the “fossilized” steady state. *Second Language Research* 14, 1-26.
- Lardiere, D. 2008. Feature-assembly in second language acquisition. In *The Role of Formal Features in Second Language Acquisition*, J. Liceras, H. Zobl & H. Goodluck (eds), 106-140. Mahwah, NJ, Erlbaum.
- Lardiere, D. 2009. Some thoughts on the contrastive analysis of features in second language acquisition. *Second Language Research* 25, 173-227.
- Larsen Freeman, D. & Long, M.H. 1991. *An Introduction to Second Language Acquisition Research*. London, Longman.
- Lee, H. 2001. Markedness and word order freezing. In *Formal and Empirical Issues in Optimality Theoretic Syntax*, P. Sells (ed.), 63-127. Stanford, CA, CSLI Publications.
- Lee, L. 2006. A study of native and nonnative speakers’ feedback and responses in Spanish-American networked collaborative interaction. In *Internet-mediated Intercultural Foreign Language Education*, J. Belz & S. Thorne (eds), 147-176. Boston, Thomson Heinle.
- Leonetti, M. 2004. Specificity and object marking: The case of Spanish *a*. *Catalan Journal of Linguistics* 3, 75-114.
- Lepschy, A.L. & Lepschy, G. 1981. *La Lingua Italiana: Storia, Varietà dell’Uso, Grammatica*. Milano, Bompiani.
- Levelt, W.J.M. 1989. *Speaking: From Intention to Articulation*. Cambridge, MA, MIT Press.
- Levelt, W.J.M., Roelofs, A. & Meyer, A.S. 1999. A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22, 1-75.
- Levin, B. & Rappaport Hovav, M. 2005. *Argument Realisation*. Cambridge, Cambridge University Press.
- Levy, M. & Stockwell, G. 2006. *CALL Dimensions: Options and Issues in Computer-assisted Language Learning*. Mahwah, NJ, Erlbaum.

- Lewis, T. & Walker, L. (eds) 2003. *Autonomous Language Learning in Tandem*. Sheffield, Academy Electronic Publications.
- Li, C.N. & Thompson, S.A. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley, CA, University of California.
- Li Wei (ed.) 2006. *The Bilingualism Reader*. London, Routledge.
- Loewen, S. 2009. Grammaticality judgment tests and the measure of implicit and explicit L2 knowledge. In *Implicit and Explicit Knowledge in Second language Learning and Teaching*. R. Ellis, S. Loewen, R. Erlam, J. Phelps, C. Elder & H. Reinders (eds), 65-93. Clevedon, Multilingual Matters.
- Long, M.H. 1996. The role of linguistic environment in second language acquisition. In *Handbook of language acquisition; Vol. 2, Second Language Acquisition*, W.C. Ritchie & T.K. Bhatia (eds), 413-468. New York, Academic Press.
- Long, M.H. 2015. *Second Language Acquisition and Task-based Language Teaching*. Chichester, West Sussex, Wiley-Blackwell.
- Lord, C. & Venter, A. 1992. Outcome and follow-up studies of High-Functioning Autistic individuals. In *High-functioning Individuals with Autism*, E. Schopler & G.B. Mesibov (eds), 187-200. New York, Plenum Press.
- Lujan, M. 1996. Clitic-doubling and the acquisition of agreement in Spanish. In *Perspectives on Spanish Linguistics*, J. Gutiérrez-Rexach & L. Silva Villar (eds), 119-138. Los Angeles, UCLA, Department of Linguistics.
- Macaro, E. (ed.) 2013. *The Bloomsbury Companion to Second Language Acquisition*. London and New York, Bloomsbury.
- Mackey, A. 1999. Input, interaction, and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition* 21, 557-587.
- Mackey, A., Pienemann, M. & Thornton, I. 1991. Rapid Profile: A second language screening procedure. *Language and Language Education* 1 (1), 61-82.
- MacWhinney, B. 1997. Second language acquisition and the Competition Model. In *Tutorials in Bilingualism: Psycholinguistic Perspectives*, A.M.B. de Groot & J. Kroll (eds), 113-142. Mahwah, NJ, Erlbaum.
- Maiden, M. 1995. *A Linguistic History of Italian*. London, Longman.
- Mansouri, F. 1995. The acquisition of Subject-Verb agreement in Arabic as a second language. *Australian Review of Applied Linguistics* 18, 65-85.
- Mansouri, F. 1997. From emergence to acquisition: Development issues in agreement marking among Australian learners of Arabic. *The Australian Review of Applied Linguistics* 20, 83-104.
- Mansouri, F. 1999. *The Acquisition of Arabic as a Second Language: From Theory to Practice*. ASLA 7. Sydney, Language Acquisition Research Centre, University of Western Sydney.
- Mansouri, F. 2002. Exploring the interface between syntax and morphology in second language development. In *Developing a Second Language: Acquisition, Processing and Pedagogy of Arabic, Chinese, English, Italian, Japanese, Swedish*, B. Di Biase (ed.), 59-72. Melbourne, Language Australia.

- Mansouri, F. 2005. Agreement morphology in Arabic as a second language: Typological features and their processing implications. In *Cross-Linguistic Aspects of Processability Theory*, M. Pienemann (ed.), 117-154. Amsterdam, Benjamins.
- Mansouri, F. & Håkansson, G. 2007. Conceptualizing intra-stage sequencing in the learner language. In *Second Language Acquisition Research: Theory Construction and Testing*, F. Mansouri (ed.), 95-117. Newcastle, UK, Cambridge Scholars.
- Matsumoto, Y. 1996. *Complex Predicates in Japanese. A Syntactic and Semantic Study of the Notion of 'Word'*. Stanford, CA, CSLI Publications, and Tokyo, Kurocio Publishers.
- Mayer, M. 1969. *Frog, Where are You?* New York, Dial Books for Young Readers.
- McLaughlin, B. 1978. The monitor model: Some methodological considerations. *Language Learning* 28, 309-332.
- McLaughlin, B. 1980. Theory and research in second language learning: An emerging paradigm. *Language Learning* 30, 331-350.
- Medojević, L. 2014. *The Effect of First Year of Schooling on Bilingual Language Acquisition: A study of Second and Third Generation Serbian-Australian 5-year-old Bilingual Children from Processability Perspective*. PhD thesis, University of Western Sydney, Sydney.
- Meisel, J. 1989. Early differentiation of languages. In *Bilingualism Across Lifespan: Aspect of Acquisition, Maturity and Loss*, K. Hyltenstam & L. Odler (eds), 13-40. Cambridge, Cambridge University Press.
- Meisel, J., Clahsen, H. & Pienemann, M. 1981. On determining developmental sequences in natural second language acquisition. *Studies in Second Language Acquisition* 3 (2), 109-135.
- Mesibov, G.B., Shea, V. & Adams, L.W. 2001. *Understanding Asperger Syndrome and High Functioning Autism*. New York, Kluwer Academic/Plenum Publishers.
- Montrul, S. 2004. *The Acquisition of Spanish. Morphosyntactic Development in Monolingual and Bilingual L1 Acquisition and Adult L2 Acquisition*. Amsterdam, Benjamins.
- Montrul, S. 2008. *Incomplete Acquisition in Bilingualism: Re-examining the Age Factor*. Amsterdam, Benjamins.
- Montrul, S. & Bowles, M. 2008. Negative evidence in instructed Heritage Language acquisition: A preliminary study of Differential Object Marking. In *Selected Proceedings of the 2007 Second Language Research Forum*, M. Bowles & al. (eds), 252-262. Somerville, MA, Cascadilla Proceeding Project.
- Montrul, S. & Bowles, M. 2009. Back to basics: Incomplete knowledge of Differential Object Marking in Spanish heritage speakers. *Bilingualism: Language and Cognition* 12 (3), 363-383.
- Montrul, S. & Sánchez-Walker, N. 2013. Differential Object Marking in child and adult Spanish Heritage Speakers. *Language Acquisition* 20 (2), 109-132.
- Müller, G. 1988. Zur Analyse subjektloser Konstruktionen in der Rektions-Bindungs-Theorie. That-trace-Effekte und pro-Lizensierung im Russischen. Seminararbeit, Frankfurt am Main, Goethe Universität.
- Mycock, L. 2007. Constituent question formation and focus: A new typological perspective. *Transactions of the Philological Society* 105 (2), 192-251.

- Næss, Å. 2007. *Prototypical Transitivity*. Amsterdam, Benjamins.
- Nespor, M. & Vogel, I. 2007. *Prosodic Phonology*. The Hague, Mouton de Gruyter.
- Nichols, J. 1986. Head-marking and dependent-marking grammar. *Language* 66, 56-119.
- Nicoladis, E., Song, J. & Marentette, P. 2012. Do young bilinguals acquire past tense morphology like monolinguals, only later? Evidence from French-English and Chinese-English bilinguals. *Applied Psycholinguistics* 33, 457-479.
- Noda, H., Sakota, K., Shibuya, K. & Kobayashi, N. 2001. *Nihongo Gakushūsha no Bunpō Shūtoku (Second Language Acquisition of Japanese Grammar)*. Tokyo, Taishukan.
- Nordlinger, R. 1998. *A Grammar of Wambaya*. Canberra, Pacific Linguistics.
- Nuzzo, E. 2012. Imparare la costruzione passiva in italiano L2. *Studi Italiani di Linguistica Teorica e Applicata* 41 (1), 67-84.
- Nuzzo, E. & Bettoni, C. 2011. Insegnamento mirato in varietà avanzate d'italiano L2: una sperimentazione didattica nel quadro della Teoria della Processabilità. *Rivista di Psicolinguistica Applicata* 11 (1-2), 33-47.
- Ohashi, J.K., Mirenda, P., Marinova-Todd, S., Hambly, C., Fombonne, E., Szatmari, P., . . . Thompson, A. 2012. Comparing early language development in monolingual- and bilingual- exposed young children with autism spectrum disorders. *Research in Autism Spectrum Disorders* 6 (2), 890-897. doi: <http://dx.doi.org/10.1016/j.rasd.2011.12.002>.
- O'Rourke, B. 2005. Form-focused Interaction in online tandem learning. *CALICO Journal* 22 (3), 433-466.
- Ortega, L. 2009. *Understanding Second Language Acquisition*. London, Arnold-Hodder Education.
- Ozonoff, S., Pennington, B.F. & Rogers, S.J. 1991. Executive function deficits in high-functioning autistic individuals: Relationship to theory of mind. *Journal of Child Psychology and Psychiatry* 32 (7), 1081-1105.
- Ozonoff, S., Young, G.S., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L., . . . Stone, W.L. 2011. Recurrence risk for Autism Spectrum Disorders: A baby siblings research consortium study. *Pediatrics* 128 (3), e488-e495. doi: 10.1542/peds.2010-2825.
- Pallotti, G. 2007. An operational definition of the emergence criterion. *Applied Linguistics* 28, 361-382.
- Paramskas, P.C. 1999. The shape of computer-mediated communication. In *Computer Assisted Language Learning (CALL): Media, Design and Applications*, K. Cameron (ed.), 13-34. Lisse, The Netherlands, Swets & Zeitlinger.
- Payne, S. & Whitney, P. 2002. Developing L2 oral proficiency through synchronous CMC: Output, working memory, and interlanguage development. *CALICO Journal* 20 (1), 7-32.
- Perlmutter, D. & Moore, J. 2002. Language-internal explanation: The distribution of Russian impersonals. *Language* 78, 619-650.
- Petersen, J., Marinova-Todd, S., & Mirenda, P. 2012. Brief report, An exploratory study of lexical skills in bilingual children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders* 42 (7), 1499-1503. doi: 10.1007/s10803-011-1366-y

- Pickering, M.J., Branigan, H.P. & McLean, J.F. 2002. Constituent structure is formulated in one stage. *Journal of Memory and Language* 46, 586-605.
- Pienemann, M. 1981. *Der Zweitspracherwerb ausländischer Arbeiterkinder*. Bonn, Bouvier.
- Pienemann, M. 1984. Psychological constraints on the teachability of languages. *Studies in Second Language Acquisition* 6, 186-214.
- Pienemann, M. 1985. Learnability and syllabus construction. In *Modelling and Assessing Second Language Acquisition*, K. Hyttenstam & M. Pienemann (eds), 23-76. Clevedon, Multilingual Matters.
- Pienemann, M. 1987. Determining the effect of instruction on L2 speech processing. *Australian Review of Applied Linguistics* 10, 83-113.
- Pienemann, M. 1989. Is language teachable? Psycholinguistic experiments and hypotheses. *Applied Linguistics* 10, 52-79.
- Pienemann, M. 1998. *Language Processing and Second Language Development: Processability Theory*. Amsterdam, Benjamins.
- Pienemann, M. 2002. The procedural skill hypothesis for SLA. In *An Integrated View of Language Development*, P. Burmeister, T. Piske & A. Rohde (eds), 43-56. Trier, Wissenschaftlicher Verlag.
- Pienemann, M. 2003. Language processing capacity. In *The Handbook of Second Language Acquisition*, C.J. Doughty & M.H. Long (eds), 679-714. Oxford, Blackwell.
- Pienemann, M. 2005. An introduction to Processability Theory. In *Cross-linguistic Aspects of Processability Theory*, M. Pienemann (ed.), 1-60. Amsterdam, Benjamins.
- Pienemann, M. 2007. Processability Theory. In *Theories in Second Language Acquisition*, B. VanPatten & J. Williams (eds), 137-154. Mahwah, NJ, Erlbaum.
- Pienemann, M. 2011a. Developmental schedules. In *Studying Processability Theory. An Introductory Textbook*. M. Pienemann & J.-U. Keßler (eds), 3-11. Amsterdam, Benjamins.
- Pienemann, M. 2011b. Explaining developmental schedules. In *Studying Processability Theory. An Introductory Textbook*. M. Pienemann & J.-U. Keßler (eds), 50-63. Amsterdam, Benjamins.
- Pienemann, M. 2011c. The psycholinguistic basis of PT. In *Studying Processability Theory. An Introductory Textbook*. M. Pienemann & J.-U. Keßler (eds), 27-49. Amsterdam, Benjamins.
- Pienemann, M., Di Biase, B. & Kawaguchi, S. 2005. Extending Processability Theory, In *Cross-linguistic Aspects of Processability Theory*, M. Pienemann (ed.), 199-251. Amsterdam, Benjamins.
- Pienemann, M., Di Biase, B., Kawaguchi, S. & Håkansson, G. 2005. Processing constraints on L1 transfer. In *Handbook of Bilingualism: Psychological Approaches*, J.F. Kroll & A.M.B. de Groot (eds), 128-153. Oxford & New York, Oxford University Press.
- Pienemann, M. & Håkansson, G. 1999. A unified approach towards the development of Swedish as L2: A processability account. *Studies in Second Language Acquisition* 21, 383-420.
- Pienemann, M. & Håkansson, G. 2007. Full transfer vs. developmentally moderated transfer. *Second Language Research* 23 (4), 485-494.

- Pienemann, M. & Johnston, M. 1986. An acquisition-based procedure for second language assessment. *Australian Review of Applied Linguistics* 9, 92-122.
- Pienemann, M. & Johnston, M. 1987. Factor influencing the development of language proficiency. In *Applying second language acquisition research*, D. Nunan (ed.), 45-141. Adelaide, National Curriculum Research Centre, Adult Migrant Education Program.
- Pienemann, M., Johnston, M. & Brindley, G. 1988. Constructing an acquisition-based procedure for second language assessment. *Studies in Second Language Acquisition* 10, 217-224.
- Pienemann, M. & Kessler, J.-U. 2007. Measuring bilingualism. In *Handbook of Applied Linguistics; Vol. 5, Multilingualism*, P. Auer & Li Wei (eds), 247-274. Berlin/New York, Mouton/de Gruyter.
- Pienemann, M. & Kessler, J.-U. 2010. *Rapid Profile: Linguistic Profiling*. <http://groups.uni-paderborn.de/rapidprofile/> (accessed 10.01.2010).
- Pienemann, M. & Kessler, J.-U. (eds) 2011. *Studying Processability Theory. An Introductory Textbook*. Amsterdam, Benjamins.
- Pienemann, M. & Kessler, J.-U. 2012. Processability Theory. In *The Routledge Handbook of Second Language Acquisition*, S.M. Gass & A. Mackey (eds), 228-246. New York, Routledge.
- Pienemann, M., Kessler, J.-U. & Lenzing, A. 2013. Developmentally Moderated Transfer and the role of L2 in L3 acquisition. In *Language Acquisition and Use in Multilingual Contexts: Theory and Practice*, A.F. Mattisson & C. Norrby (eds), 142-159. Lund, Sweden, University of Lund, Travaux de L'Institut de Linguistique De Lund.
- Pienemann, M. & Mackey, A. 1993. An empirical study of children's ESL development and rapid profile. *Studies in Second Language Acquisition* 15, 495-503.
- Pienemann, M., Mackey, A. & Thornton, I. 1991. *Rapid Profile: A Second Language Screening Procedure*. Canberra, National Languages Institute of Australia.
- Pinker, S. 1984. *Language Learnability and Language Development*. Cambridge, MA, Harvard University Press.
- Pinker, S. 2004. Clarifying the logical problem of language acquisition. *Journal of Child Language* 31, 949-953.
- Pinker, S. & Prince, A. 1988. On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition* 28, 73-193.
- Pickering, M.J., Branigan, H.P. & McLean, J.F. 2002. Constituent structure is formulated in one stage. *Journal of Memory and Language* 46, 586-605.
- Plag, I. 2008a. Creoles as interlanguage. Inflectional morphology. *Journal of Pidgin and Creole Languages* 23, 114-135.
- Plag, I. 2008b. Creoles as interlanguage. Syntactic structures. *Journal of Pidgin and Creole Languages* 23, 307-328.
- Plag, I. 2011. Pidgins and creoles. In *Studying Processability Theory. An Introductory Textbook*, M. Pienemann & J.-U. Kessler (eds), 106-120. Amsterdam, Benjamins.
- Polinsky, M. 1995. Double objects in causatives: Towards a study of coding conflict. *Studies in Language* 19, 129-221.

- Poulisse, N. 1999. *Slips of the Tongue: Speech Errors in First and Second Language Production*. Amsterdam, Benjamins.
- Qi, R. 2011. *The Bilingual Acquisition of English and Mandarin: Chinese Children in Australia*. New York, Cambria Press.
- Qi, R. & Di Biase, B. 2005. L1 and L2 patterns in the bilingual language development of a Mandarin-English child. Paper presented at the *5th International Symposium on Bilingualism*. Universitat Autònoma de Barcelona, March 20-23, 2005.
- Qi, R., Di Biase, B. & Campbell, S. 2006. The transition from nominal to pronominal person reference in the early language of a Mandarin-English bilingual child. *International Journal of Bilingualism* 10, 301-329.
- Rahkonen, M. & Håkansson, G. 2008. Production of written L2-Swedish – processability or input frequencies? In *Processability Approaches to Second Language Development and Second Language Learning*, J.-U. Kessler (ed.), 135-164. Newcastle, UK, Cambridge Scholar Publishing.
- Radford, A. 2004. *English Syntax: An Introduction*. Cambridge, UK, Cambridge University Press.
- Real Academia Española 2010. *Nueva Gramática de la Lengua Española. Manual*. Madrid, Espasa.
- Renzi, L., Salvi, G.P. & Cardinaletti, A. 2001. *Grande Grammatica Italiana di Consultazione*. Bologna, Il Mulino.
- Rizzi, L. 1997. The fine structure of the left periphery. In *Elements of Grammar: A Handbook of Generative Grammar*, L. Haegemann (ed.), 281-337. Dordrecht, Kluwer.
- Roberts, J.A., Rice, M.L. & Tager-Flusberg, H. 2004. Tense marking in children with autism. *Applied Psycholinguistics* 25 (3), 429-448.
- Robinson, P. (ed.) 2012. *Routledge Encyclopedia of Second Language Acquisition*. New York, Routledge.
- Rodriguez-Mondoñedo, M. 2008. The Acquisition of Differential Object Marking in Spanish. *Probus* 20 (1), 111-145.
- Rothman, J. & VanPatten, B. 2013. On multiplicity and mutual exclusivity: The case for different SLA theories. In *Contemporary Approaches to Second Language Acquisition*, M.P. García Mayo, M.J. Gutierrez Mangado & M. Martínez Adrián (eds), 243-256. Amsterdam, Benjamins.
- Sakai, H. 2008. An analysis of Japanese university students' oral performance in English using processability theory. *System* 36, 534-549.
- Salvi, G. & Vanelli, L. 2004. *Nuova Grammatica Italiana*. Bologna, Il Mulino.
- Sano, K. 1977. An experimental study on the acquisition of Japanese simple sentences and cleft sentences. *Descriptive and Applied Linguistics* 10, 213-233.
- Sasaki, Y. 1998. Processing and learning of Japanese double-object active and causative sentences: An error-feedback paradigm. *Journal of Psycholinguistic Research* 27 (4), 453-479.
- Savić, M. & Anđelković, D. 2007. The role of input frequency in early language production: Children's usage of Serbian prepositions. In *Frequency Effects in Language Acquisition: Defining the Limits of Frequency as an Explanatory Concept*, I. Gülzow & N. Gagarina (eds), 145-180. New York, Mouton de Gruyter.

- Schachter, J. 1974. An error in error analysis. *Language Learning* 27, 205-214.
- Schmidt, R.W. & Frota, S.N. 1986. Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In *Talking to Learn: Conversation in Second Language Acquisition*, R. Day (ed.), 237-326. Rowley, MA, Newbury House.
- Schwarz, B.D. & Sprouse, R.A. 1994. Word order and nominative case in non-native language acquisition. In *Language Acquisition Studies in Generative Grammar*, T. Hoekstra & B.D. Schwartz (eds), 317-68. Amsterdam, Benjamins.
- Schwarze, C. 2002. Representation and variation: On the development of Romance auxiliary syntax. In *Time over Matter: Diachronic perspectives in morphosyntax*, M. Butt & T. Holloway King (eds), 143-172. Stanford, CA, CSLI.
- Schwarze, Ch. 2009. *Grammatica della Lingua Italiana*. Roma, Carocci.
- Schwienhorst, K. 2002. Evaluating tandem language learning in the MOO: Discourse repair strategies in a bilingual internet project, *Computer Assisted Language Learning* 15 (2), 135-145.
- Segalowitz, N. 2003. Automaticity and second languages. In *The Handbook of Second Language Acquisition*, C.J. Doughty & M.H. Long (eds), 382-408. Oxford, Blackwell.
- Selkirk, A.G. 1984. *Phonology and Syntax. The Relation between Sound and Structure*. Cambridge, MA, MIT Press.
- Sells, P. 1995. Korean and Japanese morphology from lexical perspective. *Linguistic Inquiry* 26, 277-325.
- Sells, P. 1999. Lexicalist syntactic theories and language education. Paper presented at *AILA symposium, Waseda University, Tokyo. 5 August, 1999*.
- Sells, P. 2001. Introduction. In *Formal and Empirical Issues in Optimality Theoretical Syntax*, P. Sells (ed.), 1-16. Stanford, CA, CSLI Publications.
- Semel, E., Wiig, E.H. & Secord, W. 2003. *Clinical Evaluation of Language Fundamentals*. 4th ed. San Antonio, TX, The Psychological Corporation.
- Serianni, L. 1996. *Grammatica Italiana: Italiano Comune e Lingua Letteraria*. Milano, Garzanti.
- Seung, H., Siddiqi, S. & Elder, J.H. 2006. Intervention outcomes of a bilingual child with autism. *Journal of Medical Speech – Language Pathology* 14 (1), 53-63.
- Shibatani, M. 1987. Japanese. In *The Major Languages of East and South-East Asia*, B. Comrie (ed.), 127-170. London, Routledge.
- Shibatani, M. 1990. *The Languages of Japan*. Cambridge, Cambridge University Press.
- Shibatani, M. 1994. An integrational approach to possessor raising, Ethical datives, and adversative passives. In *Berkeley Linguistic Society* 20, 461-486.
- Shibatani, M. 2011. Japanese. In *The Major Languages of the World*, Comrie, B. (ed.), 741-764. London & New York, Routledge.
- Silva-Corvalan, C. 1994. *Language Contact and Change*. Oxford, Oxford University Press.
- Skehan, P. 1998. *A Cognitive Approach to Language Learning*. Oxford, Oxford University Press.

- Slobin, D.I. 1997. The origins of grammaticisable notions: Beyond the individual mind. In *The Crosslinguistic Study of Language Acquisition; Vol. 5, Expanding the contexts*, D.I. Slobin (ed.), 265-324. Mahwah, NJ, Earlbaum.
- Slobin, D.I. & Bever, T. 1982. Children use canonical sentence schemas: A crosslinguistic study of word order and inflections. *Cognition* 12, 229-65.
- Song, J.J. 2012. *Word Order*. Cambridge, Cambridge University Press.
- Sorace, A. 2003. Near-nativeness. In *The Handbook of Second Language Acquisition*, C.J. Doughty & M.H. Long (eds), 130-151. Oxford, Blackwell.
- Sorace, A. & Filiaci, F. 2006. Anaphora resolution in near-native speakers of Italian', *Second Language Research* 22 (3), 339-368.
- Sotillo, S. 2005. Corrective feedback via Instant Messenger Learning activities in NS-NNS and NNS-NNS dyads. *Calico Journal* 22 (3), 467-496.
- Swain, M. 1985. Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In *Input in Second Language Acquisition*, S.M. Gass & C.G. Madden (eds), 235-253. Rowley, MA, Newbury House.
- Swain, M. 1995. Three functions of output in second language learning. In *Principle and Practice in Applied Linguistics: Studies in Honor of H.G. Widdowson*, G. Cook & B. Seidlhofer (eds), 125-144. Oxford, Oxford University Press.
- Tager-Flusberg, H. 2004. Strategies for conducting research on language in autism. *Journal of Autism and Developmental Disorders* 34 (1), 75-80.
- Tager-Flusberg, H. 2005. Designing studies to investigate the relationships between genes, environments, and developmental language disorders. *Applied Psycholinguistics* 26 (1), 29-39.
- Tager-Flusberg, H. & Calkins, S. 1990. Does imitation facilitate the acquisition of grammar? Evidence from a study of autistic, Down's syndrome and normal children. *Journal of Child Language* 17 (3), 591-606.
- Tager-Flusberg, H. & Joseph, R.M. 2003. Identifying neurocognitive phenotypes in autism. *Philosophical Transactions: Biological Sciences* 358 (1430), 303-314.
- Tager-Flusberg, H., Paul, R. & Lord, C. 2005. Language and communication in autism. In *Handbook of Autism and Developmental Disorders*, F.R. Volkmar, R. Paul, A. Klin & D. Cohen (eds), 3rd ed., vol. 1, 335-364. Hoboken, NJ, Wiley & Sons.
- Tanaka, M. 2001. *The Acquisition of Point of View and Voice in Japanese as a Foreign/Second Language: The Case of Adult Learners of English, Korean, Chinese and Indonesian/Malay*. PhD thesis, International Christian University, Tokyo.
- Tankó, G. 2005. *Into Europe. The Writing Handbook*. Budapest, Teleki László Foundation.
- Taube-Shiffnorman, M. & Segalowitz, N. 2005. Within-language attention control in second language processing. *Bilingualism: Language and Cognition* 8, 195-206.
- Tek, S., Mesite, L., Fein, D., & Naigles, L. 2014. longitudinal analyses of expressive language development reveal two distinct language profiles among young children with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders* 44 (1), 75-89. doi: 10.1007/s10803-013-1853-4.

- Thomson, C.K. 2007. Gakushuu kankyoo o dezain suru: Gakushuusha komyunitihii to shite no nihongo kyooshi yoosei koosu (Designing a learning environment: Teacher training course as a learner community). *Japanese Language Education Around the Globe* 17, 169-186.
- Thorne, S. 2008. Computer-mediated communication. In *Encyclopedia of Language and Education, 2nd Edition; Vol. 4, Second and Foreign language education*, N. Van Deusen-Scholl & N.H. Hornberger (eds), 325-336. New York, Springer Science & Business Media, LLC.
- Timberlake, A. 2004. *A Reference Grammar of Russian*. Cambridge, Cambridge University Press.
- Tippets, I. 2011. Differential Object Marking: Quantitative evidence for underlying hierarchical constraints across Spanish dialects. In *Selected Proceedings of the 13th Hispanic Linguistics Symposium*, L.A. Ortiz-López (ed.), 107-117. Somerville, MA, Cascadilla Proceedings Project.
- Tomlin, R.S. 1986. *Basic Word Order: Functional Principles*. London, Croom Helm.
- Torrego, E. 1998. *The Dependencies of Objects*. Cambridge, MA, MIT Press.
- Torrego, E. 1999. El complemento directo preposicional. In *Gramática Descriptiva de la Lengua Española; Vol. 2, Las Construcciones Sintácticas Fundamentales. Relaciones Temporales, Aspectuales y Modales*, I. Bosque & V. Demonte (eds), 1779-1805. Madrid, Espasa Calpe.
- Torregrossa, J. & Bettoni, C. 2013. The development of English yes/no questions at the syntax/phonology interface. Paper presented at the 23rd Annual Conference, European Second Language Association (EUROSLA 23), Amsterdam, 28-31.2013.
- Toyoda, E. & Harrison, R. 2002. Categorization of text chat communication between learners and native speakers of Japanese. *Language Learning and Technology* 6 (1), 82-99.
- Tsujimura, N. 1996. *An Introduction to Japanese Linguistics*. Cambridge, MA, Blackwell.
- Valicenti-McDermott, M., Tarshis, N., Schouls, M., Galdston, M., Hottinger, K., Seijo, R., ... Shinnar, S. 2013. Language differences between monolingual English and bilingual English-Spanish young children with Autism Spectrum Disorders. *Journal of Child Neurology* 28(7), 945-948. doi: 10.1177/0883073812453204
- Vallduvi, E. & Engdahl, E. 1996. The linguistic realisation of information packaging. *Linguistics* 34, 459-519.
- Van Valin, R.D. 2001. *Introduction to Syntax*. Cambridge, Cambridge University Press.
- Van Valin, R.D. 2005. *Exploring the Syntax-Semantics Interface*. Cambridge, Cambridge University Press.
- VanPatten, B. 1996. *Input Processing and Grammar Instruction: Theory and Research*. Norwood, NJ, Ablex.
- VanPatten, B. 2007. Input processing in adult second language acquisition. In *Theories in Second Language Acquisition*, B. VanPatten & J. Williams (eds), 115-135. Mahwah, NJ, Erlbaum.
- VanPatten, B. & Cadierno, T. 1993. Explicit instruction and input processing. *Studies in Second Language Acquisition* 15, 225-241.

- VanPatten, B. & Williams, J. (eds) 2007. *Theories of Second Language Acquisition*. Mahwah, NJ, Erlbaum.
- Vermeulen, R. 2009. On the syntactic typology of topic marking: A comparative study of Japanese and Korean. *UCL Working Papers in Linguistics* 21, 335-36.
- Vigliocco, G., Butterworth, B. & Garrett, M.F. 1996. Subject-verb agreement in Spanish and English: Differences in the role of conceptual constraints. *Cognition* 61, 261-298.
- Vigliocco, G., Butterworth, B. & Semenza, C. 1995. Constructing subject-verb agreement in speech: The role of semantic and morphological factors. *Journal of Memory and Language* 34, 186-215.
- Vigliocco, G. & Franck, J. 1999. When sex and syntax go hand in hand: Gender agreement in language production. *Journal of Memory and Language* 40, 455-478.
- Vigliocco, G. & Franck, J. 2001. When sex hits syntax: Gender agreement in sentence production. *Journal of Memory and Language* 45, 368-390.
- Vigliocco, G., Hartsuiker, R.J., Jarema, G. & Kolk, H.H.J. 1996. One or more labels on the bottles? Notional concord in Dutch and French. *Language and Cognitive Processes* 11 (4), 407-442.
- Vihman, M.M., & McLaughlin, B. 1982. Bilingualism and second language acquisition in preschool children. In *Verbal Processes in Children: Progress in Cognitive Development Research*, C.J. Brainerd & M. Pressley (eds), 35-58. New York, Springer.
- Vincent, N. 2011. Italian. In *The Major Languages of Eastern Europe*, B. Comrie (ed.), 233-252. London, Routledge.
- Volkmar, F.R., Reichow, B., Westphal, A., & Mandall, D.S. 2014. Autism and the Autism Spectrum: Diagnostic concepts. In *Handbook of Autism and Pervasive Developmental Disorders*, F.R. Volkmar, S.J. Rogers, P. Rhea & K.A. Pelphrey (eds), 4th ed., vol. 1, 3-27. Hoboken, NJ, Wiley & Sons.
- Wang, K. 2006. *The Acquisition of the English Passive Construction by Chinese Learners of ESL*. Unpublished BA (Honours) thesis, University of Western Sydney, Sydney.
- Wang, K. 2009. Acquiring the passive voice: Online production of the English Passive construction by Mandarin speakers. In *Research in Second Language Acquisition: Empirical Evidence Across Languages*, D. Keatinge & J.-U. Kefler (eds), 93-117. Newcastle, UK, Cambridge Scholars Publishing.
- Wang, K. 2010. *The Acquisition of English Passive Constructions by Mandarin Speakers: A Developmental Perspective*. Unpublished PhD thesis, University of Western Sydney, Sydney.
- Wang, K. 2011. 'The Fish was Eaten': *The Acquisition of the English Passive Construction by Chinese Learners of ESL*. Koln, LAP Lambert Academic Publishing.
- Ware, P. & Cañado, M. 2007. Grammar and feedback: Turning to language form in telecollaboration. In *Online Intercultural Exchange: An Introduction for Foreign Language teachers*, R. O'Dowd (ed.), 107-126. Clevedon, UK, Multilingual Matters.
- Ware, P. & O'Dowd, R. 2008. Peer feedback on language form in telecollaboration. *Language Learning & Technology* 12 (1), 43-63.
- Warschauer, M. 1996. Comparing face-to-face and electronic discussion in the second lan-

- guage classroom. *CALICO Journal* 13 (2), 7-26.
- Waterhouse, L. & Fein, D. 1982. Language skills in developmentally disabled children. *Brain and Language* 15 (2), 307-333.
- Weyerts, H., Penke, M., Münte, T.F., Heinze, H.-J. & Clahsen, H. 2002. Word order in sentence processing: An experimental study of verb placement in German. *Journal of Psycholinguistic Research* 31, 211-268.
- WHO 1993. *The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research*. Geneva, World Health Organization.
- White, L. 1989. *Universal Grammar and Second Language Acquisition*. Amsterdam, Benjamins.
- Wilson, F. 2009. Processing at the syntax-discourse interface in second language acquisition. PhD Thesis. University of Edinburgh.
- Wirbatz, K. 2014. *Investigating the Fundamental Difference between L1 and L2 Acquisition Based on Syntax: A Fresh Look at the Development of German L1 and L2*. PhD thesis, University of Western Sydney, Sydney.
- Yamaguchi, Y. 2008. The early syntactic development in child L2 acquisition: What happens after "canonical order"? In *Processability Approaches to Second Language Development and Second Language Learnings*, J.-U. Keßler (ed.), 245-266. Newcastle, UK, Cambridge Scholars Publishing.
- Yamaguchi, Y. 2009. The development of plural marking and plural agreement in child English L2 acquisition. In *Research in Second Language Acquisition: Empirical Evidence Across Languages*, D. Keatinge & J.-U. Keßler (eds), 9-39. Newcastle, UK, Cambridge Scholars Publishing.
- Yamaguchi, Y. 2010. *The Acquisition of English as a Second Language by a Japanese Primary School Child: A Longitudinal Study from a Processability Viewpoint*. PhD thesis, University of Western Sydney, Sydney.
- Yamaguchi, Y. 2013. *Child Second Language Learning: A Study of English as a Second Language*. Saarbrücken, LAP Lambert Academic Publishing.
- Yamaguchi, Y. & Kawaguchi, S. 2014. Acquisition of English morphology by a Japanese school-aged child: A longitudinal study. *Asian EFL Journal* 16 (1), 89-119.
- Yip, V. & Matthews, S. 2007. *The Bilingual Child: Early Development and Language Contact*. New York, Cambridge University Press.
- Zhang, Y. 2002. A processability approach to L2 acquisition of Chinese grammatical morphemes. In *Developing a Second Language: Acquisition, Processing and Pedagogy of Arabic, Chinese, English, Italian, Japanese, Swedish*, B. Di Biase (ed.), 29-44. Melbourne, Language Australia.
- Zhang, Y. 2004. Processing constraints, categorical analysis, and second language acquisition of the Chinese adjective suffix *de* (ADJ). *Language Learning* 54, 437-468.
- Zhang, Y. 2005. Processing and formal instruction in the L2 acquisition of five Chinese grammatical morphemes. In *Cross-Linguistic Aspects of Processability Theory*, M. Pienemann (ed.), 155-177. Amsterdam, Benjamins.
- Zhang, Y. 2007. Testing the Topic Hypothesis: The L2 acquisition of Chinese syntax. In *Second Language Acquisition Research: Theory Construction and Testing*, F. Mansouri (ed.), 145-172. Newcastle, UK, Cambridge Scholars.

- Zhang, Y. & Widyastuti, I. 2010. Acquisition of L2 English morphology: A family case study. *Australian Review of Applied Linguistics* 33 (3), 29.1–29.17. doi: 10.2104/aryl1029.
- Zobl, H. 1986. Word order typology, lexical government and the prediction of multiple, graded effects in L2 word order. *Language Learning* 36, 159-183.
- Zwaigenbaum, L., Bryson, S., Rogers, T., Roberts, W., Brian, J. & Szatmari, P. 2005. Behavioral manifestations of autism in the first year of life. *International Journal of Developmental Neuroscience* 23 (2-3), 143-152.

About the authors

Tonya G. Agostini is a PhD candidate at the MARCS Institute at the University of Western Sydney, Australia. Her current research programme investigates early acquisition of Italian L2 in children with high functioning autism spectrum disorder. Her research interests include early word learning and morphophonological development in L1 and L2 acquisition in typical and atypical populations, such as those with autism spectrum disorder, specific language impairment and dyslexia.

Daniele Artoni is a PhD student at the University of Verona, where he also tutors in Russian literature. His research interests are second language acquisition, with a focus on grammatical development in Russian. His wider interests include literature and music.

Catherine T. Best is Professor and Chair of Psycholinguistic Research at the University of Western Sydney, Australia, and has a long-standing affiliation as a senior scientist at Haskins Laboratories, USA. She received her PhD in developmental psychology and human neuropsychology, and gained subsequent training in phonetics and phonology. She held prior academic appointments at Columbia University and Wesleyan University (USA). Her research focuses on how language experience shapes cross-language speech perception in infants and adults, and influences adults' and young children's recognition of spoken words across regional accents of their native language.

Camilla Bettoni is Professor at the University of Verona, Italy. A graduate from the Universities of Padua and Edinburgh, she held teaching appointments in Scotland, Australia, and Italy, and visiting positions in Germany (Konstanz), Sweden (Stockholm), and Austria (Klagenfurt). Her research interests include second language grammar development; first language loss; cross-cultural pragmatics; and the (socio)linguistics of migration, with particular attention to Italian, as acquired by immigrants in Italy and lost by emigrants abroad in contact with English.

Bruno Di Biase, is Associate Professor at the School of Humanities and Communication Studies and at MARCS Institute, University of Western Sydney, Australia. A close collaborator with Manfred Pienemann on many projects since the establishment of the Language Acquisition Research Centre (LARC) at the University of Sydney in the early 1990s, he received his PhD in linguistics at the Australian National University. His research focuses on second language development and the applications of SLA research to language teaching and learning in formal and natural language environments, including situations of migration and language contact.

Giorgia Ginelli (BA, Verona 2003; MA, Padova 2005; PhD, Verona 2010) wrote her doctoral thesis on the acquisition of constituent questions in Italian L2. A teacher and teacher trainer, she is chairperson of *Italiano con noi*, a school of Italian language and culture in Verona. Her interests include language teaching methodology, especially the deployment of drama and theater techniques in the language classroom.

Barbara Hinger is Professor of language teaching methodology (*Fachdidaktik*) at the University of Innsbruck, Austria. Her research areas are second language acquisition and classroom-based language assessment, with a focus on Spanish L2. She is chairwoman of ÖGSD (*Österreichischen Gesellschaft für Sprachendidaktik*, the Austrian Association of Language Pedagogy), and a member of the board of VERBAL (*Verband für Angewandte Linguistik*, the Austrian affiliate of the International Association of Applied Linguistics).

Louise Jansen is a Lecturer in Applied Linguistics and German Studies at the Australian National University. Her primary research interest is second language development, with a focus on Processability Theory. Her recent data-based work includes the acquisition of L2 German syntax in instructed settings (Jansen 2008) and nominal plural (Charters, Dao & Jansen 2011).

Satomi Kawaguchi is a Senior Lecturer in Japanese, second language acquisition, and TESOL research methods at the University of Western Sydney. Her research areas include Japanese L2 and English L2. She is interested in theory-construction in SLA and research-driven-teaching, as well as in the use of new communication technology and student-focused innovative pedagogies. She was awarded the Australian Teaching & Learning Council Citation for outstanding contribution to student learning in 2010.

Marco Magnani is a PhD student at the University of Verona, where he also tutors undergraduate students in Russian language and literature. His research interests include second language acquisition, with a focus on grammatical development in Russian and in Italian.

Lucija Medojević is an applied linguist at the School of Humanities and Communication Arts, University of Western Sydney, Australia. She holds a BA Hons in Languages & Linguistics (University Medal) and a PhD from the University of Western Sydney. Her research interests include bilingual/first/second language acquisition, the development of narrative skills in bilingual children, and the connection of biculturalism to creative writing.

Yumiko Yamaguchi, PhD, is an associate professor at the Foreign Language Center, Tokai University, Japan. She obtained master's degree in Education from Boston University, USA, and a PhD in Applied Linguistics from the University of Western Sydney, Australia. Her current research focuses on ESL/EFL acquisition and language teaching methodology.