

Constructing a second language inventory –  
the accumulation of linguistic resources in L2 English.

Ph.D. dissertation  
Institute of Language and Communication  
University of Southern Denmark

Supervisor: Johannes Wagner, University of Southern Denmark  
Co-supervisor: Carl Bache, University of Southern Denmark

Søren Wind Eskildsen  
Institute of Language and Communications  
University of Southern Denmark  
DK-5230 Odense M  
Denmark

2008



Constructing a second language inventory –  
the accumulation of linguistic resources in L2 English.

**Contents.**

<b>People and Places: Acknowledgments</b>	<b>i</b>
<b>Introduction</b>	<b>1</b>
Preliminaries	1
The research interest	2
The data	4
About the anthology	5
<b>Chapter 1:</b>	<b>10</b>
<b>EC:</b> Søren W. Eskildsen and Teresa Cadierno: Are recurring multi-word expressions really syntactic freezes? Second Language Acquisition from the perspective of Usage-Based Linguistics.	
<b>Chapter 2:</b>	<b>23</b>
<b>ESK1:</b> Constructing Another Language – Usage-Based Linguistics in Second Language Acquisition.	
<b>Chapter 3:</b>	<b>50</b>
<b>ESK2:</b> What’s New? - Routines and Creativity along a Usage-Based Path of Second Language Learning	
<b>Chapter 4:</b>	<b>87</b>
<b>ESK3:</b> Second Language Learning as Participation and Acquisition: Towards a new SLA Eclecticism.	
<b>Chapter 5:</b>	<b>127</b>
<b>ESK4:</b> You no oich – User-Based L2 Learning: The Case of Negation	
<b>Chapter 6: Portable language experience – a brief outline of UBL</b>	<b>169</b>
6.1 Introduction	165
6.2 UBL: Maximalism and other basic tenets	165
6.3 UBL vs. UG: logical problems of language learning?	170

6.4 UBL for SLA: setting the stage	175
<b>Chapter 7: Linguistic resources and routines – MWEs and other units in L2 learning</b>	<b>181</b>
7.1 Introduction	181
7.2 Setting the stage	181
7.3 On starting points: Linguistic Units and L2 Vocabulary Research.	183
7.4 Idioms, Open Choices, and Collocations	188
7.5 On the right terms	193
7.6 Formulaic language – findings in SLA	204
7.7 Perspectives – formulaic language, UBL, and SLA	209
<b>Chapter 8: E Pluribus Unum? – tracing the roots of SLA</b>	<b>212</b>
8.1 Introduction	212
8.2 There's competence and then there's competence – a brief Introduction to SLA's adolescence	212
8.2.1 Interlanguage and psycholinguistics	213
8.2.2 Developmental sequences	217
8.2.3 Communicative Competence	224
8.3 Individual competence or social performance?	228
8.3.1. Is there a super-theory of SLA?	229
8.3.2 Tracing the social-cognitive split	239
8.3.3 Second language acquisition or second language participation?	249
8.4 Overcoming the dualisms	261
<b>Chapter 9: Conclusions and perspectives</b>	<b>265</b>
Complete References	272
Appendices: Inventory counts and traced utterances, initial and final recording periods (ESK2)	296
Appendix 1: Valerio, recording period 1	296
Appendix 2: Valerio, recording period 6	303
Appendix 3: Carlos, recording period 1	315
Appendix 4: Carlos, recording period 4	321
English Summary	339
Dansk Resumé	345

## **People and Places: Acknowledgments**

A confession: This dissertation is a piece of work which I have not carried out entirely by myself. I certainly did the writing; I know I did all the transcribing. But I did not do all the thinking on my own. Thus, I am greatly indebted to people who took part in the thinking process. They are, in a form of chronological order, Rineke Brouwer who first assisted me in the process of putting together my dissertation proposal. She is the first person without whom I could not have done this. The second person must be Teresa Cadierno; her name as co-author on my first published work says it all. The third person – and the one who was with me all the way – was my supervisor, Johs Wagner. Thank you, Johs, for inspiring me to keep doing what I do. Without Rineke, Teresa, and Johs there certainly would have been no dissertation from my hand.

Other people from other places have been helpful, too. I am especially indebted to Steve Reder, Kathy Harris, and John Hellermann from the Applied Linguistics Department at Portland State University for inviting me to Portland and for granting me access to their unique database – and for teaching me how to use the software needed to access the data. I also want to thank the 'tech team' at the Lab School in Portland, Vid and Glen, who would always provide technological first aid when it was needed. I also owe a great deal of thanks to both teachers and research assistants at the Lab School for letting me share their office space when in Portland.

Of course, I must also thank my own local spot in Odense, Institut for Sprog og Kommunikation, for approving my dissertation proposal and giving me a work space. I am especially indebted to my colleagues in the Language and Cognition Research Group for inspirational discussions. Also thanks to my co-supervisor, Carl Bache, who was always available when needed.

I dedicate this work to those people mentioned here; they are the ones without whom I could not have done this. To you, Lise, I can only say I'm sorry – times will change now as I get out of my hole. You're not the reason why I did this work. You're the reason why I look forward to getting out of the hole where, it must have appeared to the outside world, I've been stuck for three years. This work is not for you; everything else is.

## **Constructing a second language inventory – the accumulation of linguistic resources in L2 English.**

### **Introduction.**

#### Preliminaries.

When people learn languages they do so in many different ways and by many different means. They are cognitively active; they figure out the target language as they engage in it. They are interactionally active; they establish conversational routines as they take part in it. They are socially competent people; they are people doing something real in the real world. The real world of the present research is a second language (L2) classroom in Portland, Oregon. The real people, the focal students of the research, are Mexican-Spanish speakers. They go to class, they sit at their desks, they perform the classroom activities they are asked to perform, they interact with their teachers and their classmates. Therefore they are many different identities; Mexican immigrants, appliance repair men, married, single – but they are also bona fide L2 learners of English. The question is, what, in this real world, do they learn as they engage in the L2 classroom interaction?

The present research reflects a general increased interest, within Second Language Acquisition (SLA) research, in what might be broadly referred to as 'the mental lexicon'. I do not, from the onset, wish to imply that I take a special interest in 'words' over other aspects of language and language acquisition, but I do wish to imply that 'grammar' does not really do it for me – as I believe it does not do it for my focal students. Grammar, as traditionally conceived, is not something L2 learners learn, neither *in* the real world, nor to *deploy in* the real world. The basic concern of the present research is to investigate the recurring building blocks of language as they are made relevant in the context of learning a L2, the primary interest focusing on empirical investigations of what might be said to be the kernel of L2 linguistic knowledge; the units of language learning.

Such an undertaking requires a well-thought-out linguistic theory, an elaborate view of learning, as well as an idea of what it means to take part in social interaction. Contrary to much traditional SLA, where the underlying linguistic theory is usually left unmentioned but where the object to be learned is usually defined in terms of 'words and rules', where learning is never discussed but

unfailingly accepted a priori to be about the structured acquisition of 'input', operationalised as 'intake' to denote the idea that not everything L2 learners hear around them is internalised, the present research aims to study the learning of recurring bits and pieces of language as they are used in interaction. To this end, I invoke Usage-Based Linguistics (UBL) and its notion that all linguistic units are form-meaning pairings implying that semantically empty combinatorial rules of language are not relevant in terms of L2 learning. I discuss metaphors of learning, as participation and acquisition, to denote the idea that something as complex as L2 ontogenesis cannot be captured by recourse to one such metaphor only – and to imply that, in terms of participation, learning happens, not in an interactional vacuum, but in a real world where interaction matters as more than a site for information exchange and negotiation for meaning.

#### The research interest.

The present research can therefore be said to bring together two strands of SLA – or perhaps more precisely, aspects of these strands, these being 1) the focus on 'lexical units' (though not 'words', as traditionally thought of); and 2) the recent increased tendency to combine cognitively and socially oriented approaches to the study of language learning (Larsen-Freeman 2007). I perform this operation by calling upon UBL which insists on defining the linguistic unit in specific terms and always with recourse to meaning and function, and by attempting to contextualize L2 development in terms of local language emergence. In other words, I view the object of research in L2 learning as the accumulation of interactional resources and routines; i.e., I investigate linguistic patterns and how they develop over time. This makes for a strictly empirical operationalization of linguistic knowledge which demands an abolition of the strict dichotomy between competence and performance. The only a priori decision to be made is the decision to take UBL as my point of departure.

This lack of predetermined linguistic structures to look for stems from an insight generated by the interest in looking for recurrent 'formulaic sequences' (Wray 2002) in the data. Initially, the present research was framed around an attempt to investigate the role of such formulaic sequences in L2 learning, an area of SLA research still not blessed with unequivocal results (Schmitt and Carter 2004). However, it quickly transpired that such formulaic sequences, partly because of their largely elusive nature in linguistic theory and research, seemed to a large degree to be absent from the data. The question of whether L2 learners start out from formulas and gradually start analysing them to

use the individual constituents in other linguistic structures, or if they start from the learning of combinatorial rules with practice ensuring the entrenchment of certain formulas over time, seems fundamentally misguided and flawed. Formulas and more general patterns seem to co-exist at all points in development, at least as far as the data investigated here are concerned. One does not happen before the other in ontogenesis; there is no formulaic language ahead of current interlanguage competence – and there is no current interlanguage competence ahead of formulaic language. Such convictions, it will be argued, stem from what Langacker termed 'the rule-list fallacy'.

Instead, the research focus became one of investigating the extent to which the learning of all such 'formulas' and general patterns could be said to be item-based along a usage-based path of learning from fixed patterns via partially flexible patterns to completely schematic constructions, as suggested in UBL. As such the investigations revolve around the notion of L2 learning as the gradual accumulation of an assortment of interactional routines and resources. Combined with this linguistic descriptive and analytical tool-box from UBL, the concept of emergentism, at whose heart is the notion that language knowledge is fundamentally in flux, will be invoked to throw light on the seemingly essentially constant and never-ending nature of L2 learning

Hence, the research as presented in the five research papers is informed by the idea that the field of SLA lacks a framework within which to study an empirically based, performance-driven, emergent linguistic inventory of a L2 learner. The strictly field-related epistemological interest lies in charting such emergent inventories for the first time. The studies presented in the five research papers in this collection are therefore exploratory in nature, seeking out the fabric of those linguistic inventories. Only linguistic units actually found in the data are discussed; no units, either of a lexically specific or a syntactically generic nature, are invoked to explain linguistic development. Rather, what I see in my data is always and everywhere described, analysed, and explained with recourse to the linguistic and interactional reality of the learners. This is done on the basis of the UBL framework and its insistence on using actual language use in actual usage events (Langacker 2000) as its point of departure for doing research.



### The data.

The data source for all the research papers is the Multimedia Adult English Learner Corpus (MAELC) which consists of audio-visual recordings of classroom interaction in an English as a Second Language (ESL) classroom in Portland, Oregon. MAELC was compiled and is maintained at The National Labsite for Adult ESOL (known locally as the Lab School<sup>1</sup>). The Lab School was a partnership between Portland State University and Portland Community College. The ESL classrooms, in which the recordings were made, were equipped with six ceiling-mounted video cameras. Four of those were fixed, and two were moveable by remote control. The two latter cameras each followed a student wearing a wireless microphone; students were given these microphones to wear on a rotational basis. The teacher wore a microphone at all times in the class (Reder et al. 2003; Reder 2005).

The final database of the inquiries in the five research papers consists of transcripts from approx. 70 classroom sessions each consisting of three hours of recordings in which my focal students are either wearing a microphone or sitting next to someone wearing a microphone. Carlos, the focal student in EC, ESK1, ESK2, and ESK4, attended ESL class from September 2001 through February 2005, and Valerio, the focal student in ESK2, ESK3, and ESK4, attended class from July 2003 through July 2005. During their time in class, both Valerio and Carlos were considered to be successful learners as they gradually progressed, by the standards of this language program, from Level A to Level D (beginning to intermediate; for more information on the proficiency levels, see Brillanceau 2005; Reder 2005).

Taking as its point of departure those two Mexican-Spanish speaking classroom learners of English, the present research investigates L2 learning as a constant movement of linguistic advancement – but a movement which has no visible endpoint (Firth and Wagner 1998), a movement where completion is always deferred (Hopper, quoted in Lantolf and Thorne 2006: 14). The emergent linguistic inventories of the two focal students are described and analysed in terms of an interactionally situated grammar consisting of recurring flexibly abstract units of actual language use. Language learning in this sense is dealt with in terms of linguistic and interactional progression

---

<sup>1</sup> The Lab School was supported, in part, by grant R309B6002 from the Institute for Education Science, U.S. Dept. of Education, to the National Center for the Study of Adult Learning and Literacy (NCSALL). I am deeply indebted to Steve Reder and all the staff at the Lab School for granting me access to their data, providing me with office space while in Portland to work on the data, and generally being extremely hospitable and helpful.

and discussed in the light of the two aforementioned learning metaphors. The investigation of linguistic inventories along these lines is done in the form of the five research papers; their methodological and theoretical underpinnings and perspectives are further unpacked and discussed in the subsequent chapters. The rest of this introduction serves as a guide to this collection of research papers and theoretical-methodological chapters.

#### About the anthology:

Starting with the five research papers, these investigate various aspects of emergent linguistic inventories along a usage-based path of learning, going from formulas through low-scope patterns to fully schematised abstract constructional language knowledge (e.g., Tomasello 2000, 2003; N. Ellis 2002). In this anthology they are presented in the first five separate chapters. Chapter 1, Eskildsen and Cadierno (EC)<sup>2</sup>, studied negation development in one of the focal students, suggesting that the usage-based path of learning was indeed a valid default for investigating the acquisition of L2 structure. Seen retrospectively, it formed the preamble for the subsequent research as presented in the four remaining papers, ESK1, ESK2, ESK3, and ESK4 respectively. In terms of the chronology of those four papers, it is also evident that there is a movement towards an increased awareness of the importance of studying interactional development in the classroom alongside the more traditional linguistic inquiries. Chapter 2, ESK1<sup>3</sup>, thus showed the fundamentally locally situated nature of the initially occurring 'formulas' in learning, the items from which the linguistic inventories are seen to spring, operationalised as recurring multi-word expressions (MWEs). They were operationalised as such to denote the usage-based vantage point both in terms of the non-distinction between lexis and syntax as separate compartments of language, and in terms of underlining the methodological principle that no a priori structures were defined as formulas; these were allowed to rise from the data.

Chapters 4 and 5, ESK3<sup>4</sup> and ESK4<sup>5</sup>, expand on these insights. Thus, echoing findings from language socialization studies in L2 learning (e.g., Kanagy 1999; Hellermann 2006), showing how

---

<sup>2</sup> Published in 2007. *Collocations and Idioms 1: Papers from the First Nordic Conference on Syntactic Freezes*, Joensuu, May 19-20, 2007. Editors M. Nenonen and S. Niemi. Studies in Languages, University of Joensuu, Vol. 41. Joensuu: Joensuu University Press

<sup>3</sup> Under review for publication in *Applied Linguistics*.

<sup>4</sup> Under review for publication in *International Review of Applied Linguistics*.

<sup>5</sup> Under review for publication in *L2 Learning as Social Practice: Conversation-analytic Perspectives* (Working title). Editors G. Pallotti and J. Wagner.

interactional settings fertilize locally contextualized linguistic growth, ESK 3 shows how routinized activities co-develop with linguistic routines – but it also showed (and herein lies the expansion) that linguistic routines may spill over into new social contexts, rendering some linguistic routines transcendent in nature. ESK4, on the other hand, taking a slightly different investigative road, investigates the coming together of interactional and linguistic paths. The primacy of this study was to confirm the findings in EC by applying the investigative modus operandi to another focal student; but in order to study in more detail the individual differences which emerged between the two learners and to examine the emergent nature of a specific, frequent occurrence of a specific learner pattern, the study showed the ontogenetical fusion between the focal student's general interactional routines and capabilities and the development of his linguistic inventory.

The data analysed in ESK3 and ESK4 shows the need for an elaborate investigative framework to tackle the interactional phenomena encountered. It does not seem sufficient to call on the notion of usage event as it is found in UBL. Therefore, some micro-analytic tools, mostly inspired by Conversation Analysis, are applied to account for local interactional contingencies found in the data. This results in an eclectic approach to the study of developmental issues in L2 learning, suggesting the need to abolish or rethink certain dualisms, among them being the performance-competence distinction and the dichotomous relationship between the two conflicting notions of learning as captured in the 'participation' and 'acquisition' metaphors, respectively (Sfard 1998).

Chapter 3, ESK2, is situated in the middle of this. It investigates the empirical totality of the two linguistic inventories and concludes that the two focal students throughout development operate on a motley assortment of linguistic patterns. These patterns, not surprisingly, were found to include MWEs – but more surprisingly, they were found to be overwhelmingly represented by a different kind of patterns, namely 'utterance schemas'. Taken from Tomasello (2000), utterance schemas are partially open, partially closed (i.e., lexically filled) constructions. Such patterns, e.g., *I can verb* or *how many x*, are highly frequent, and it seems that in terms of ontogenesis their emergence is dependent on previously used patterns. ESK2, then, viewed emergent linguistic creativity as the discovery of new patterns, each closely related to previous ones, as well the discovery of an increasing number of items to put in the open slots in the utterance schemas.

In sum, the five research papers investigate emergent linguistic inventories in terms of an empirically valid ever-evolving grammar consisting of real units recurring in use. The new contribution in relation to existing research into L2 linguistic development is this search for an empirically grounded recurring linguistic unit; a common denominator of recurring patterning as it is done and carried out by people learning a L2. The insights generated by the incipient recognition that L2 learning – possibly also when described and analysed in purely linguistic terms – never happens in a social vacuum have resulted in an approach that seeks to situate individual L2 learners interactionally and frame issues of learning along two metaphorical dimensions, namely learning as participation and learning as acquisition. As such, the present research attempts to instigate a new way of doing L2 learning research; a new form of SLA eclecticism. This operation involves an exploration of why 'existing units' do not suffice, why the prevailing notion of formulaic language is inadequate, and what UBL can do to improve the state of affairs along those lines. It also involves an exploration of what separates the present approach from existing ones in SLA. The chapters following the five research papers will provide these discussions.

Chapter 6 outlines the UBL framework. It is not intended as an exhaustive account of UBL – such an account lies far beyond the scope of the present dissertation, considering the existence within that framework of rather extensive linguistic theories, including construction grammar in various guises (e.g., Goldberg 1995; Croft 2001; Croft and Cruse 2004), Langacker's cognitive grammar (Langacker 1987, 1991; Taylor 2002), and Hopper's emergent grammar (e.g., Hopper 1987, 1998), as well as a range of elaborate child language acquisition studies (e.g., Tomasello 1992, 2000, 2003; Dabrowska 2000; Dabrowska and Lieven 2005). Rather, the point of the chapter is to elaborate on the UBL conception of linguistic knowledge as a usage-derived collection of linguistic expressions and constructions, which, by way of the importance programmatically ascribed to fixed and semi-fixed pieces of language, is much better suited for dealing with the role of formulaic language in SLA and its interplay with more productive language in development.

Chapter 7, on linguistic resources and routines, is concerned with the notion of the linguistic unit as it has traditionally been characterised in SLA research. It contains a brief outline of L2 vocabulary studies and the conceptualisation of the linguistic unit usually fertilized here, but the central aspect of Chapter 1 is the discussion of formulaic language, how it is traditionally conceived mostly in a psycholinguistic vein based on a compartmentalization of linguistic knowledge into what is defined by syntactic rules and what is stored in the mental lexicon. In fact, as will be shown in chapter 7 and

8, this view seems to dominate not only research in formulaic language and L2 vocabulary acquisition, but also SLA research as it has been carried out traditionally.

Chapter 8 contextualizes my 'UBL for SLA' approach in a more general discussion of the field of SLA. Tracing the roots of SLA and adding to an ongoing discussion concerning 'the social vs. the individual' within the research field, this chapter will position this approach in relation to traditional mainstream psycholinguistic approaches to SLA, most prominently the Input/Interactionist framework (Gass 2003), as well as in relation to the socio-cognitive approach (Atkinson 2002; Atkinson et al. 2007), the socio-cultural approach (e.g., Lantolf 2000, 2005), and 'CA for SLA' (e.g., Brouwer and Wagner 2004; Markee and Kasper 2004; Hellermann 2007). Drawing from the research results in the five research papers Chapter 8 will make the case for a socially oriented UBL as the presently most profitable framework within which to investigate linguistic development in a L2.

I would like to stress the individuality of the five research papers at this point. Presented in chapters 1-5 they are distinct and independent, albeit interrelated, investigations. As such they stand on their own. They are not chapters that coherently and naturally follow each other like chapters in a book – for this is not a book, it is an anthology of research papers held together by the use of the same database as well as aspects of methodology, approach, and linguistic theory. This has bearings on the lay-out of the dissertation. The research papers are printed here in their original forms as published / submitted which means that their respective original formats are retained. They are each complete with their own references and appendices (ESK2 and ESK3). Footnotes are also kept as in the original article manuscripts; however, in terms of formatting, the footnote numeration runs throughout the dissertation. The appendices at the end of the dissertation are extra material to back the methodology and findings in ESK2 – because of their size, they were not submitted to the journal along with the manuscript. The complete references towards the end of the dissertation include all references, from research papers and background chapters alike, in order to provide the reader with an overview of discussed literature.

The aspects holding the research papers together include the conviction that UBL is a psycholinguistically valid model of language knowledge and a growing recognition based on accruing empirical evidence, that language use and learning cannot fruitfully be kept apart.

Together, these two assumptions make for an approach to SLA research that represents a new SLA eclecticism, related to and building on generated insights from child language acquisition, especially Tomasello (2003), and a diversity of related L2 research (e.g., Firth and Wagner 1997; Larsen-Freeman 1997; Atkinson 2002; N. Ellis and Larsen-Freeman 2006; de Bot et al. 2007; Hellermann 2007). In proposing this eclecticism I am not claiming to answer all questions in SLA research; I will not have much to say about the co-constructed nature of much interaction (Learner 1991), and there are social issues of identity construction and relationship building in the community of practice (Brouwer and Wagner 2004; Hellermann 2006) about which I do not have much to offer. What I am claiming is to present a framework which enables linguistically and interactionally interested researchers to work in a coherent fashion with incipient language knowledge as it takes shape in the world of L2 classroom interaction; a framework which allows researchers to study individual linguistic development against the backdrop of a developing interactional competence (e.g., Kramsch 1986; Hall 1993; Young 2000; Hellermann 2007).

## Chapter 1: EC.

*Søren W. Eskildsen and Teresa Cadierno*  
*University of Southern Denmark*

### Are recurring multi-word expressions really syntactic freezes? Second Language Acquisition from the perspective of Usage-Based Linguistics.\*<sup>6</sup>

#### 1. Introduction:

This study focuses on the development in the oral use of negation patterns by a classroom Mexican learner of English. In theoretical agreement with recent claims made in various areas of second language acquisition (SLA) research, e.g., matters of frequency (Ellis 2002), learner constructions (Waara 2004), motion constructions (Cadierno 2004, Cadierno and Ruiz 2006), the aim of the study is to discuss the role of *recurring multi-word expressions* (MWEs) in L2 acquisition and use from the perspective of Usage-Based Linguistics (UBL). For our purposes UBL is particularly relevant as theoretical framework because it acknowledges the importance of MWEs and because it does not dichotomise syntax and lexis. This, we argue, allows for a better understanding of the structure, meaning, use, and acquisition of MWEs. In this paper we present our theoretical approach in more detail and present some results of our on-going investigations in SLA.

#### 2. Theoretical underpinnings:

According to Tummers *et al.* (2005: 228) UBL comprises a number of related linguistic theories which are united by three main principles. These principles are “the priority assigned to language use, the integration of competence and performance, and the rejection of the rule-list fallacy.” The first principle implies that usage is the only valid source for investigating the systemic nature of language. Language structure is, in other words, seen to emerge from concrete usage situations, and “it is the task of the linguist to study the whole range of repetition in discourse and to seek out those regularities that promise interest as incipient subsystems” (Hopper 1998: 166). In terms of the

---

\* Published in Nenonen, M and S. Niemi (eds.). 2007. *Collocations and Idioms 1: Papers from the First Nordic Conference on Syntactic Freezes, Joensuu, May 19-20, 2007*. Studies in Languages, University of Joensuu, Vol. 41. Joensuu: Joensuu University Press. The material presented in this paper is part of two ongoing research projects conducted by the two authors on the use of multi-word expressions by non-native speakers of English. One project, which is financed by a Ph.D.-grant at the University of Southern Denmark, takes ESL classroom interaction as its empirical point of departure. The other project, which is partly financed by the Danish Council of Humanities, focuses on the use of multi-word expressions in globalised organisational contexts, where English is used as a lingua franca.

<sup>6</sup> The authors would like to thank John Hellermann for useful comments on an earlier version of the paper.

individual language user, this means that usage mirrors knowledge and that linguistic knowledge is conceptualised as linguistic experience (Tomasello 2000). It also implies a link between constructions as conventionalised in the speech community and as entrenched in the mind of the individual (Ellis 2002). This belief that language use and language knowledge, interaction and cognition, individuality and sociality are mutually constitutive is directly linked with the second UBL principle that competence and performance are seen as interwoven rather than dichotomous.

With respect to the third principle mentioned above, UBL rejects the view of language knowledge as dichotomised along the lines of 1) something that is rule-governed, cognitively demanding, and flexible and 2) something that is listed in the mental lexicon, swiftly processed, and stable. Related to the UBL conceptualisation of language knowledge as a structured inventory of symbolic units, i.e., conventionalised form-meaning pairings, used for communicative purposes (Langacker 1987), this is the rejection of the rule-list fallacy and its “[...] assumption, on grounds of simplicity, that particular statements (i.e., lists) must be excised from the grammar of a language if general statements (i.e., rules) that subsume them can be established (Langacker 1987: 29).

Adult language knowledge, then, is seen to consist of a continuum of linguistic constructions of different levels of complexity and abstraction comprising “concrete and particular items (as in words and idioms), more abstract classes of items (as in word classes and abstract constructions), or complex combinations of concrete and abstract pieces of language (such as mixed constructions)” (Tomasello 2003: 99). It should be kept in mind, however, that only *adult* language knowledge has this systemic nature. The emergent nature of language structure as described above means that ontogenetically, children do not learn their L1 on the basis of an innate Universal Grammar. Rather, operating “with different psycholinguistic units than adults” (Tomasello 2000: 62), children learn language in an item-based fashion heavily reliant on frequency, recurrence, and imitation. Only when enough exemplars (items) have been encountered by the child does she begin to build the abstract cognitive schemas thought to underlie language knowledge. This view of language learning as item-based combined with the rejection of the rule/list fallacy means that storage and learning may take place on multiple levels simultaneously, involving both concrete items and abstract schemas, which may cohabitate in the grammar (Achard 2006). For example, storage for PLURAL may be dually represented as a) car + s = cars; and b) THING + [morph] = more things.



In opposition to the UBL principles outlined above, previous research into Formulaic Language (Wray's (2002) term; FL) has been predominantly 'syntactocentric', influenced by formalistic, generative views on language. This influence is clear in Wray (2002: 143) where it is stated that "language knowledge entails developing rules to generate all the possible utterances of the language". Dominated by the rules/list fallacy, this has led to the view that whatever is 'formulaic' is somehow inferior to syntactic rules (van Lancker-Sidtis and Rallon 2004). Applying the UBL framework, thereby challenging the strict division between creative expressions generated by syntax, on the one hand, and lexicon-based expressions on the other, we aim to challenge this. Whereas Wray (2002) finds it premature to implement UBL as a theoretical framework in research in FL, we argue that because of its attempts to account for all kinds of usage patterns without compartmentalising language, it allows for a better understanding of the role of MWEs in language usage and language learning.

### 3. Formulaic language in second language acquisition (SLA) research:

With the publication of books by Wray (2002) and Nesselhauf (2005), the anthology edited by Schmitt (2004) as well as conferences such as *Collocations and Idioms: The First Nordic Conference on Syntactic Freezes*, FL has come of age as a field of research. Even though van Lancker-Sidtis & Rallon (2004: 211) describe the field as "handicapped by a bewildering array of variously defined terms", it is possible, in terms of SLA, to find some general trends in the existing research in FL. Without going into detail (but see e.g. Gitsaki 1999, Nesselhauf 2005, Wray 2002 for full discussions), we note that most research concerns collocational competence, Verb+NP, or Adv-Adj collocations, and is carried out with written language. Results are mixed in some respects, but the following conclusions are generally reached (Nesselhauf 2005): a) collocational production presents problems for learners, and more serious problems than general vocabulary use; b) learners use fewer collocations than native speakers; c) learners are not aware of neither collocational restrictions nor combinatorial potential of lexical items.

Intriguingly, even though deviant collocational usage by L2 learners is sometimes reported to be as low as 25% (Nesselhauf 2005), statistics are nonetheless read as confirming the status of collocational competence as problematic for L2 learners. This, it is argued here, is a result of the underlying view FL as deviant from the norm of grammatically generated language. It is even implied in Wray (2002: 196) that "learning formulaic language is not 'real' language learning".

Real language learning, in this view, presupposes analysability, combinability, and computation based on rules of syntax. In an inherently circular mode of argumentation, FL is seen as evidence that it is formulaic because it suggests inconsistencies between FL and grammatically combinatorial knowledge; learners are seen to produce language that is simply beyond their current interlanguage competence. In other words, FL is considered formulaic, i.e. *frozen*, because the underlying competence system, for some reason, could not have produced it – and this is so because it is formulaic.

From the perspective of UBL, the distinction presented above between FL and rule-governed syntax is a manifestation of the rule-list fallacy applied to SLA. Instead of making it a question of either FL or rules, UBL more conveniently accounts for the role of FL in language usage and language learning because it allows for dual representation, cf. above. In Tomasello's (2003: 106) words: "[...] in usage-based approaches a given linguistic structure may exist psychologically for the speaker both as a concrete expression of its own [...] and at the same time, as an exemplar of some more abstract construction [...] The main point from an acquisition point of view is that when a higher abstraction is made the lower level concrete constructions and expressions do not necessarily go away but remain available for use – especially if they are used frequently."

This introduces the notions of type and token frequency and their proposed importance for language acquisition. Token frequency refers to the idea that frequency of a concrete expression "in the language learner's experience tends to entrench that expression in terms of the concrete words and morphemes involved", whereas type frequency "of a class of expressions determines the abstractness or schematicity of the resulting construction" (Tomasello 2003: 107). In other words, storage as wholes is dependent on token frequency, whereas schematised knowledge and therefore productivity is dependent on type frequency (see also Ellis 2002).

With those aspects of frequency in mind, we set out to investigate for SLA the validity of the following proposed path of child language acquisition going from concrete formulas via low-scope patterns (part concrete, part abstract) to abstract constructions (Ellis 2002, Tomasello 2000, 2003). Formulas are holophrases, single words, and so-called frozen phrases, sometimes idiosyncratic, which convey "a holistic, undifferentiated communicative intention, most often the same communicative intention as that of the adult expressions from which they were learned."

(Tomasello 2003: 36). Examples are *lemme see, I wanna do it*. Low-scope patterns include pivot schemas, i.e., still concrete pieces of language, with one item structuring the utterance or determining the speech act, such as *where's the X, It's a X, X gone*, as well as verb-island constructions. These are utterance schemas revolving around specific verbs and constructions in which each verb is an island in the sense that morphological and syntactic markings are first learned on a verb-by-verb basis and not immediately generalised to other verbs. The final point on the proposed path concerns abstract constructions, i.e., higher-level constructional schemas that cut across different verbs, such as transitive constructions, negation constructions etc. The present study thus seeks to examine whether or not this “L1 acquisition sequence [...] could serve well as a reasonable default in guiding the investigation in which exemplars and their type and token frequencies determine the second language acquisition of structure (Ellis 2002: 170).<sup>7</sup>

#### 4. The study:

##### 4.1. Design:

Our source of data is the Multimedia Adult English Learner Corpus (MAELC)<sup>8</sup> which consists of audio-visual recordings of classroom interaction in an English as a Second Language (ESL) context (Reder *et al.* 2003). The classrooms in which the recordings took place were equipped with video cameras and students were given wireless microphones on a rotational basis; the teacher also wore a microphone. Our study is a longitudinal case study consisting of recordings from September 2001 through February 2005. The informant is an adult male learner of English, his native language is Spanish (he is from Mexico), and he became a US resident approx. 21 months prior to taking ESL classes. Our data, then, contrary to traditional research in second language collocational knowledge, cf. above, consist of oral classroom interaction. In our view, given the fact that FL is thought to enhance on-line language processing, verbal interaction makes for more intuitively appealing data than written language. The study has limits, however, in that we only have access to classroom interactions. The data presented here will only pertain to that context. Whatever goes on in the

---

<sup>7</sup> We also acknowledge, with Ellis (2002), the differences between L1 and L2 acquisition and the ways in which L2 learners differ from L1 learners, including conceptual development, use of conscious problem-solving and deduction abilities, access to input and general context of learning, and L1 → L2 transfer issues.

<sup>8</sup> MAELC is maintained and developed under the direction of Steve Reder and Kathryn Harris at The National Labsite for Adult ESOL (known locally as the Lab School). The Lab School is supported, in part, by grant R309B6002 from the Institute for Education Science, U.S. Dept. of Education, to the National Center for the Study of Adult Learning and Literacy (NCSALL). The Lab School is a partnership between Portland State University and Portland Community College.

world outside the classroom has no bearing on this study, nor can we say anything about our informant's language use outside the classroom.

#### 4.2. Analysis:

We made queries into the MAELC database and found several students that would potentially qualify as informants in longitudinal research. We filtered the queries and ended up with approx. 40 sessions (each consisting of three hours of recordings) in which our informant is either wearing a microphone or sitting next to someone wearing a microphone. We then transcribed the recorded interactions and the resulting transcripts form the final database of our enquiry. In the process, we left out private speech (Ohta 2001) as well as interactions where the students were clearly practicing certain forms that included the negation pattern. We did some unmotivated looking in the transcripts, and we quickly realised that negated contexts would be a good place to start our analysis because negated constructions seemed to be present throughout our informant's development, and because negation is a relatively transparent form-meaning pattern. In alignment with the general research dicta of UBL, our analysis, then, is based on distributional analyses of samples of language during and across particular developmental periods (Tomasello 2000).

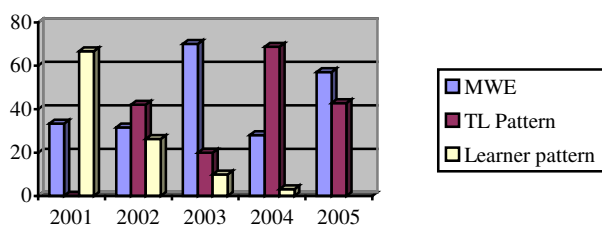
By operationalising FL as *recurring multi-word expressions* we attempt to make it clear that we are not working on a definition of FL based on a distinction between listed information in the lexicon and abstract rules of syntax. Rather, a MWE is a sequence of words used together for a relatively coherent communicative purpose. This definition is based on Tomasello's (2003) definition of 'construction', but whereas Tomasello, when defining construction, refers to an abstract level of language knowledge, we focus on concrete recurrences of specific words. In terms of analysing general recurrence in learner production, i.e. output frequencies, we distinguish between type and token frequency in order to capture the recurrences of both concrete expressions and more abstract patterns and constructions. In this we reflect the UBL conceptualisation of language knowledge as a continuum of linguistic constructions, cf. section 2 above. Note also at this point that we assume that questions of frequency pertain to output as well as input. This assumption is based on the UBL principles, mentioned in section 2 above, that language usage and language knowledge as well as social interaction and individual cognition, are mutually constitutive, and that there is a link between what is conventionalised in the speech community and entrenched in the kind of the individual.

## 5. Results:

### 5.1. Negation pattern usage development:

Fig. 1 below presents an overview of the negation patterns found in the data and how their usage frequencies develop over time. Initially, we observed three distinct but related kinds of negation pattern in the data, a) recurrent target language (TL) MWE: *I don't know*; b) learner pattern: *Subj no V* (e.g., *I no remember*); c) TL pattern: *aux-neg* pattern (e.g., *I don't think so*). In 2001, negations are divided between 1/3 MWE and 2/3 learner patterns. In 2005, there is a 42-58% division between MWE and other instances of the TL pattern. In between we have varying stages of competition between the learner pattern and the TL pattern.<sup>9</sup>

Figure 1: Negation pattern usage development



We note three tendencies: stability and importance of the MWE, increase in TL pattern usage, and decrease (until the point of disappearance) in learner pattern usage. This gives us a rough developmental path, in terms of the language inventory, on which we tentatively pinpoint three chronologically overlapping phases:

- 1: TL MWE *I don't know* + Learner pattern
- 2: TL MWE *I don't know* + Learner pattern + TL pattern
- 3: TL MWE *I don't know* + TL pattern

These tentative results imply that going further into the data, we would expect to see a fairly high and stable token frequency of the MWE and an increasingly higher type frequency in the TL pattern

<sup>9</sup> Keeping in mind Tomasello's idea of children's psycholinguistic units being different from those of adults, we note that the same might be the case for L2 learners vs. native speakers in that our informant for some time employs a distinct 'learner pattern'.

usage across time. This would then suggest the entrenchment and whole-unit status of the MWE and the development in increasingly more abstract knowledge in terms of the TL pattern. It would also lend credence to the cohabitation in the grammar hypothesis.

### 5.2. Type and token frequencies:

Table 1 below presents type and token frequencies as well as type-token ratios for the TL patterns and the learner pattern across time. Type frequency indicates variation in negated verbs, but it does not take into account variance in SUBJ-position; the higher the type-token ratio, the more varied and productive the usage of the pattern. In 2001, though, only one verb is negated (*know*), thus there is only one type and by implication the type-token ratio is incalculable. Although we cannot say anything definite about issues of storage, the token frequency of the TL pattern for this period is relatively high, supporting the whole-unit status of *I don't know*. This corresponds to what we expected for the relation between token frequency and MWEs. The development in type-token ratio suggests that the TL pattern is becoming increasingly more abstract in its schematic representation as productivity is expanded in terms of negated verbs. A system is emerging and the learner pattern seems to disappear.<sup>10</sup> Table 1 thus supports what we had expected.

*Table 1: Type and token frequencies for the TL pattern and the learner pattern*

	TL Pattern			Learner Pattern		
	Token	Type	Ratio	Token	Type	Ratio
2001	5	1	--	10	8	0,80
2002	14	7	0,50	5	4	0,80
2003	9	3	0,33	1	1	--
2004	31	13	0,42	1	1	--
2005	14	7	0,50	0	0	0

<sup>10</sup> Please note that we are not postulating an endpoint of acquisition neither in this learner nor generally speaking. We cannot be certain that things disappear from linguistic experience. For example, one instantiation of the learner pattern in 2002 seems to be a strategic self-repair by our informant, in a situation where, using the TL pattern, he is repeatedly not understood by a fellow student. This implies that the learner pattern may be a useful interactional resource and it underlines the need for future research on how aspects of interaction, e.g., turn taking and turn constructional units in conversation (e.g, Selting 2000), interplays with the learning of the constructions of a second language.

Two things are puzzling, however. For instance, 2004 sees the highest amount of different negated verbs without yielding the highest type-token ratio and therefore not showing the highest level of creativity, and 2002 sees the same type-token ratio as 2005, seemingly suggesting the same level of abstractness for those two years and by implication decrease in abstractness in 2003-4. First, in the case of the level of creativity in 2004, we found that the informant not only used the MWE *I don't know* quite frequently, but other possible MWE candidates as well, namely, *I don't think so*, *I don't remember*, and *I don't have NP*. *I don't have NP* and *I don't remember* were not counted as MWEs in this study because they were only used in 2004. *I don't think so* comes closer to qualifying as a MWE in that it is employed by this learner from an early point in development; however, we chose not to include it as MWE because its token frequency is markedly lower than that of *I don't know*.<sup>11</sup> Thus, in 2004 the TL pattern was most frequently used with four recurring verbs, explaining the lower type-token ration for that period. What we cannot explain statistically is the productivity in 2002 and the idea that the informant's language inventory should be more abstract at this point than later on in development. The relatively high numbers in 2002 could be due to a number of reasons that promise interest as focal points of future research, e.g., the nature of tasks in classroom, and functional requirements in interaction. However, we may be able to explain the phenomenon by examining more closely the existing variance in the negation patterns.

### 5.3. Degree of abstractness in negation patterns:

Table 2 below presents the degree of abstractness in negation patterns, and how it develops across time. It shows what negation patterns are employed at what stages and with what pronouns and what tense morphology. The figure supports what was suggested by the type-token ratios above, namely, that the learner is working on an increasingly abstract language system. We see this in the expansion of the combinatorial possibilities. Consequently, the explanation for the high productivity / creativity in 2002 seems to be that the learner is working on a pattern which is not as abstract in its schematicity as the construction that is emerging but abstract enough to be productive. It resembles what Ellis (2002) calls a "low-scope pattern" and could be likened with what Tomasello (2003) refers to as a pivot schema. The difference between the 'pattern' and the 'construction' is thus a matter of abstractness as alluded to earlier.

*Table 2: Negation pattern abstractness development*

---

<sup>11</sup> In total the type-token ratio of *I don't think so* is 0,07; for *I don't know* it is 0,44.

	2001	2002	2003	2004	2005
1 <sup>st</sup> pers., sing.	TL / LP	TL	TL / LP	TL / LP	TL
2 <sup>nd</sup> pers., sing.	LP	TL / LP			TL
3 <sup>rd</sup> pers., sing.	LP	LP			TL
1 <sup>st</sup> pers., pl.				TL	
2 <sup>nd</sup> pers., pl.					
3 <sup>rd</sup> pers., pl.				TL	
Past tense				TL	TL

*Note: TL=Target Language Pattern; LP=LearnerPattern; pers.= person; sing.= singular; pl=plural.*

The three phases on the developmental path that we tentatively pinpointed earlier have emerged as robust and are reiterated below; this time, in terms of more elaborate constructional representation:

1: **TL MWE:** I don't know + **Learner pattern:** PRN<sup>sing</sup> neg V<sup>prs</sup>

2: **TL MWE:** I don't know + **Learner pattern:** PRN<sup>sing</sup> neg V<sup>prs</sup> + **TL pattern:** PRN<sup>sng</sup>;  
1st,2nd aux-*do*<sup>prs</sup> neg V

3: **TL MWE:** I don't know + **TL construction** PRN aux-*do*<sup>prs/pst</sup> neg V (S)<sup>12</sup>

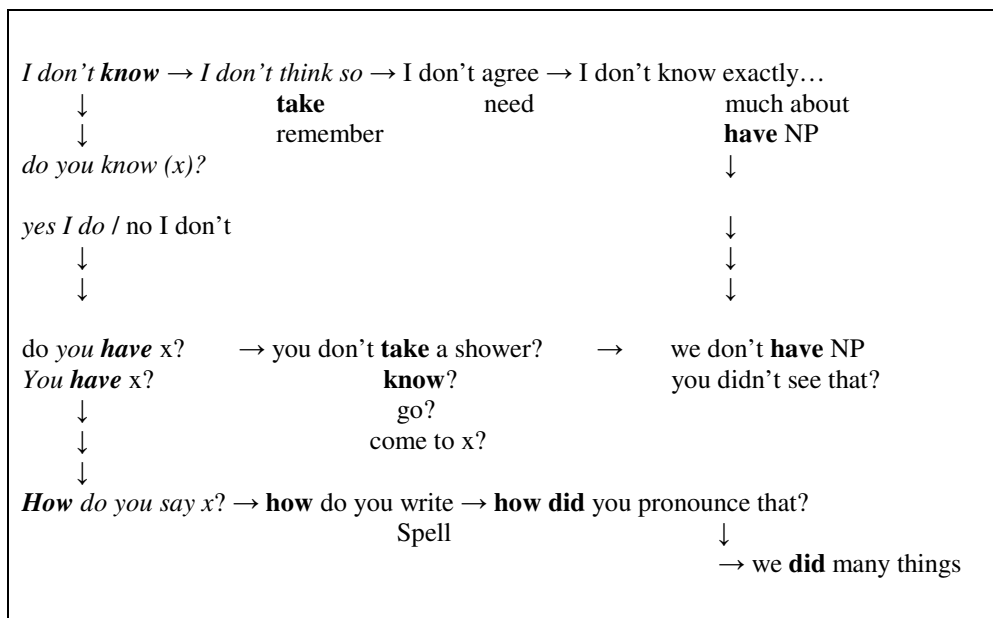
This representation makes the increasing abstractness clear, and one notes with interest that the learner pattern is just as abstract in its representation as the TL pattern; in fact, in combinatorial terms it seems to be more productive, with the TL pattern only including 1<sup>st</sup> and 2<sup>nd</sup> person pronouns as subject. The emergence of the TL construction coincides with a more varied use of pronouns and, importantly in terms of abstract schematicity, it enables the learner to express past tense negation. Table 3 below is a sketchy attempt to capture this fluid development from the concrete item to the abstract system, and the relationship of the *aux-neg* construction with some other significant uses of *aux do*, showing how the negation construction seems to be used and learnt, by our informant, in a network of interrelated emergent constructions. Development is

<sup>12</sup> Arguably, the representation should not contain the 'prs/pst' notation; the 'aux' category should stand on its own, involving this possibility implicitly. We have included it to emphasise the difference in underlying representation between the learner pattern and the learner construction.



depicted horizontally, phases vertically. Repeated lexical items (intra- as well as inter- constructionally) are in bold types, MWE candidates in italics:

Table 3: Overview of the inventory: The Aux-do network



While depicting the development of the *Aux-do network*, Table 3 also reveals the need for further research in (at least) the following areas: a) usages of *can't* and *won't* in relation to the *do*-schema; b) *copula*-negation and *have*-negation in relation to *do*-negation; c) past tense usage of the *do*-schema in relation to other usages of past tense; d) development of the *PRN* in the *do*-schema in relation to other pronoun usages; e) the role of other MWEs in other constructions; and f) other usages of the negated verbs.

## 6. Conclusions:

The analyses presented above indicate that from the MWE *I don't know* emerge increasingly abstract patterns which sanction systematic usage expansion of the negation pattern to include other lexical items, i.e. varied usage of PRNs and main verbs, and, perhaps more importantly in terms of level of abstractness, past tense expression. The system that is seen to emerge in this fashion is the gradual abstraction of regularities that link expressions as constructions. Second language learning may thus be seen as exemplar-based (Ellis 2002).

Furthermore, our data support two key UBL principles. The first relates to notions of frequency; productivity of a given constructional schema seems to be dependent on type frequency, and entrenchment of a given MWE seems to be dependent on token frequency. In our data, the instantiations of the *aux-neg* schema got more varied across time, whereas the expression *I don't know* was retained throughout development as a MWE. This supports our assumption that questions of frequency pertain to output as well as input. The second principle is that language knowledge, acquisition, and storage may be dually represented, cf. the rule/list fallacy. In our data *I don't know* as a MWE is stored as a whole that, in turn, becomes sanctioned by the very schema whose emergence it initiates.

Finally, our study does not tally well with the idea of MWEs as syntactic freezes, as a MWE in our conceptualisation is not seen as a representation of underlying syntactic rules and not necessarily schematised by underlying abstract language knowledge. Furthermore, the elements of the MWE *I don't know* cannot be thought of as frozen in this pattern as they are employed elsewhere by the learner; initially, though, the negation element is not. This does not mean, however, that the ability to negate linguistic material is frozen in this one initial instantiation of the *aux-neg* construction; remember that the learner is perfectly capable of negating in general; initially, he merely uses another pattern to do so productively, namely the learner pattern *subj no V*. Nothing is frozen here; rather, the MWE forms the backbone of schematic development, and across time it becomes a reflection of the emergent abstract language knowledge.

### References:

Achard, Michel. 2006. Usage-based semantics: Meaning and distribution of three French 'breaking' verbs. Paper presented at Collocations and Idioms: The First Nordic Conference on Syntactic Freezes, Joensuu 19-20/5-06.

Cadierno, Teresa. 2004. Expressing motion events in a second language: A cognitive typological approach. *Cognitive linguistics, second language acquisition and foreign language pedagogy*, ed. by Michel Achard & Susanne Niemeier, 13-49. Berlin: Mouton de Gruyter.

Cadierno, Teresa and Lucas Ruiz. 2006. Motion events in Spanish L2 acquisition. *Annual Review of Cognitive Linguistics* 4. 183–216.

Ellis, Nick C. 2002. Frequency effects in language processing – a review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition* 24. 143-188

- Gitsaki, Christina. 1999. *Second language lexical acquisition. A Study of the Development of Collocational Knowledge*. San Francisco: International Scholars Publications.
- Hopper, Paul J. 1998. Emergent grammar. *The new psychology of language*, volume 1, ed. by Michael Tomasello, 155-175. Mahwah, NJ.: Lawrence Erlbaum
- Langacker, Ronald W. 1987. *Cognitive grammar*, vol. 1. Theoretical prerequisites. Stanford, CA: Stanford University Press.
- Nesselhauf, Nadja. 2005. *Collocations in a learner corpus*. Amsterdam: John Benjamins.
- Ohta, Amy S. 2001. *Second language acquisition processes in the classroom*. Mahwah, NJ.: Lawrence Erlbaum
- Reder, Stephen, Kathryn A. Harris, and Kristen Setzler. 2003. A multimedia adult learner corpus. *TESOL Quarterly* 37, 3. 546-557.
- Schmitt, Norbert. (ed.) 2004. *Formulaic sequences*. Amsterdam: John Benjamins.
- Selting, Margret. 2000. The construction of units in conversational talk. *Language in Society* 29. 477-517
- Tomasello, Michael. 2000. First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics* 11-1/2. 61-82
- Tomasello, Michael 2003. *Constructing a Language*. Cambridge: Cambridge University Press.
- Tummers, Jose, Kris Heylen, and Dirk Geeraerts. 2005. Usage-based approaches in Cognitive Linguistics: A technical state of the art. *Corpus Linguistics and Linguistic Theory* 1-2. 225-261
- Van Lancker-Sidtis, Diane and Gail Rallon. 2004. Tracking the incidence of formulaic expressions in everyday speech: Methods for classification and verification. *Language and Communication* 24. 207-240.
- Waara, Renee. 2004. Construal, convention, and constructions in L2 Speech. *Cognitive linguistics, second language acquisition and foreign language pedagogy*, ed. by Michel Achard and Susanne Niemeier, 51-75. Berlin: Mouton de Gruyter.
- Wray, Alison. 2002. *Formulaic language and the lexicon*. New York, NY: Cambridge University Press.

## Chapter 2: ESK1.

### Constructing another Language – Usage-Based Linguistics in Second Language Acquisition\*

#### **Abstract.**

The general aim of this paper is to discuss the application of Usage-Based Linguistics (UBL) to an investigation of developmental issues in second language acquisition (SLA). Particularly, the aim is to discuss the relevance for SLA of the UBL suggestion that language learning is item-based, going from formulas via low-scope patterns to fully abstract constructions. It is investigated how well this suggested path of acquisition serves 'as a default in guiding the investigation of the ways in which exemplars and their type and token frequencies determine the second language acquisition of structure.' (N. Ellis 2002: 170). As such, it builds on and further discusses the findings in Bardovi-Harlig (2002) and Author and Colleague (2007). The empirical point of departure is longitudinal oral second language classroom interaction and the focal point is the use of *can* by one student in the class in question. The data reveal the formulas, here operationalised as recurring multiword expressions, to be situated in recurring usage events, suggesting the need for a fine-tuning of the UBL theory for the purposes of SLA research towards a more locally contextualized theory of language acquisition and use. The data also suggest that semi-fixed linguistic patterns, here operationalized as utterance schemas, deserve a prominent place in L2 developmental research.

#### **1. Introduction**

A number of related linguistic theories fly under the banner of UBL (e.g., Barlow and Kemmer 2000; Tummers et al. 2005). A core principle uniting these theories is the rejection of the syntax-lexicon dichotomy and an embrace of a 'maximalistic language model knowledge in which abstract grammatical patterns and the lexical instantiations of those patterns are jointly included, and which may consist of many different levels of schematic abstraction' (Tummers et al. 2005: 228-9). Language knowledge is seen as a structured inventory of symbolic units, i.e., form-meaning patterns (Langacker 1987), of varying complexity (from morphemes to full utterances) and abstractness (from fixed, perhaps idiosyncratic, concrete formulas to abstract schematic templates which in turn sanction the single instantiations).

---

\* This article is under review for publication in *Applied Linguistics*. This is a revised version of a paper presented at the Second International Conference of the German Cognitive Linguistics Association, Munich, October 6<sup>th</sup>, 2006. I would like to thank Teresa Cadierno, Gabriele Kasper, Johannes Wagner, and three anonymous reviewers for constructive comments on earlier versions of the paper. Needless to say, I am responsible for any remaining flaws.

The most important contribution of UBL to applied linguistics is the work on child language acquisition. UBL argues that children must learn 'two faces of grammar: smaller elements and larger patterns' (Tomasello 2003: 41), as they construct their first language (L1) on the basis of concrete exemplar material, encountered in real-time linguistic interaction. Rejecting the argument from the poverty of the stimulus (Tomasello 2003; MacWhinney 2004), UBL views language acquisition as experiential and item-based and directed towards increased cognitive schematicity, going from formulas via low-scope patterns to fully abstract constructions (e.g., Dabrowska 2000; Tomasello 2003; Dabrowska and Lieven 2005). This paper applies this learning trajectory to the ontogenetic construction of a L2; however, a few notes on the terminology might be in order.

Notoriously difficult to capture terminologically, formulas are most often defined as rote-learned chunks (e.g., *Lemme-see, I wanna do it, I dunno*). The term used here is recurring multi-word expression (MWE), which will be operationalized in the next section. Low-scope patterns (e.g., *I wanna Verb, I don't Verb*) consist of a fixed part and an open slot. In this paper, the term utterance schema, borrowed from Tomasello (2000), will be used to refer to patterns that are partially fixed and partially productive as some schematic knowledge is seen to sanction the use of given lexical items in the open slot. The posited level of ultimate abstractness consists of schematic knowledge of symbolic units; e.g., *I don't Verb* is cataloged more generally in the inventory as the negation construction *NP aux neg Verb*. In terms of constructing their language, children thus move from the specific expression (*I dunno*) to the general schematic construction (*NP aux neg Verb*) as regularities are abstracted, linking expressions as constructions.

The schematic generalities that are extracted from recurring patterns in ontogenesis all derive from the speaker's biographical database of experience (Tomasello 2000; N. Ellis 2002). Subject to slight change every time a new utterance is encountered (N. Ellis 2002; Tummers et al. 2005), linguistic knowledge as experience in this sense is emergent, permeable and flexible, constantly under construction in linguistic interaction and therefore in flux as environments change (Hopper 1998). This emergent individual linguistic inventory is rooted in 'usage events' (Barlow and Kemmer 2000; Langacker 2000) which describe the totality of discourse aspects, including e.g., social and cultural practices, in which all linguistic utterances are rooted. Token and type frequencies are then thought to determine matters of entrenchment and schematicity of recurring expressions and

constructions. In Tomasello's (2003: 107) words, token frequency is frequency of a concrete expression which 'in the language learner's experience tends to entrench that expression in terms of the concrete words and morphemes involved', whereas type frequency 'of a class of expressions determines the abstractness or schematicity of the resulting construction'. The maximalistic nature of the linguistic inventory, however, guarantees that what is learned as specific is not necessarily replaced over time by the more abstract constructional knowledge acquired; rather, abstract patterns and their specific instantiations may cohabitate in the grammar (Achard 2007).

UBL finds kinship in both past and present linguistic research. Bolinger (1979) and Peters (1983), giving prominence to memory over a combinatorial system, argued in favour of multiple storage of lexical items in a range of patterns in a manner which retrospectively seems to have heralded the arrival of the maximalism of UBL. Furthermore, branches of corpus linguistics share UBL's experiential view of language knowledge (Hoey 2007<sup>13</sup>) as well as its rejection of the syntax-lexis dichotomy and the ensuing view of differences among language patterns as matters of abstraction (e.g. Stefanowitsch and Gries 2003; Hoey 2007; Stubbs 2007).

## **2. Previous research**

This paper is related to three strands of research. These concern, respectively, formulaic language (FL) in SLA, *can*-usage in SLA, and UBL in SLA. Space does not permit me to discuss any of these in detail, but I encourage the interested reader to consult referenced work.

### **2.1 FL**

Since scholars such as Bolinger (1979), Pawley and Syder (1983), and Peters (1983) pointed out that people are dependent on memorized chunks, in communication as well as language learning, FL has been investigated from a variety of perspectives and under a variety of terms. In fact, some 40 terms (Wray and Perkins 2000) have been applied to capture the phenomenon, resulting in a terminologically handicapped field (van Lancker-Sidtis and Rallon 2004). Many scholars today, however, accept the definition of formulas as chunks that are stored and accessed as wholes and not generated or analyzable by the grammar (e.g. Wray and Perkins 2000; Wood 2002; Wray 2002; Butler 2003; Read and Nation 2004; Schmitt and Carter 2004; van Lancker-Sidtis and Rallon 2004; Bardovi-Harlig 2006). This psycholinguistic equivalent to Sinclair's (1991) famous division of

---

<sup>13</sup> I thank an anonymous reviewer for pointing out this similarity between UBL and Hoey's work.

corpus linguistic data into 'the idiom principle' and 'the open choice principle' is based on a compartmentalization of linguistic knowledge into 1) something that is listed in the mental lexicon, swiftly processed, and stable; and 2) something that is rule-governed, cognitively demanding, and flexible (Author and Colleague 2007). Such a stance, however, is fundamentally incommensurable with the principles of UBL, according to which all utterances are drawn from the same cognitive resource – a 'construct-i-con' (Goldberg 2003). From this perspective, place and manner of storage cannot be decisive criteria for identifying FL. Rather, formulas and productive schematic templates 'occupy opposite ends of the continuum of linguistic structures' (Achard 2007: 1). To highlight the programmatic non-division of syntax and lexis and to underscore the usage-driven methodology, this investigation adopts from Author and Colleague (2007) the operationalization of the MWE as a *recurring sequence of words used together for a relatively coherent communicative purpose*.

Research into the role of FL in L2 learning has not yielded systematic results (Schmitt and Carter 2004). In keeping with a compartmentalized view of language knowledge, however, there seems to be consensus that FL deviates from the learner's combinatorial interlanguage system; either FL is seen to be beyond the learner's interlanguage (e.g., Bolander 1989; Nattinger and DeCarrico 1992; Myles et al. 1999; Bardovi-Harlig 2002) or lag behind it (e.g., Irujo; Nesselhauf 2005)<sup>14</sup>. Thus, some scholars have argued that FL is a performance phenomenon with no links to an independently developing grammar (e.g., Krashen and Scarcella 1978; Bohn 1986; Granger 1998), whereas others agree that FL is developmentally significant and feeds into the rest of the system in development (e.g. Hakuta 1974; Nattinger and DeCarrico 1992; Weinert 1995; Myles et al. 1998; 1999; Schmitt and Carter 2000; Bardovi-Harlig 2002; Wood 2002).

Both positions underline the primary concern in FL research in SLA, namely the relationship between FL and creative grammar. Accordingly, some scholars have advocated a view of language knowledge as a formulaic-creative continuum (e.g. Nattinger and DeCarrico 1992; Weinert 1995). The maximalistic model of language knowledge envisioned by UBL entails precisely such a view. The UBL framework with its perceived fluidity among linguistic patterns and its default assumption that language learning is fundamentally a matter of abstracting generalities from recurring utterances, is particularly suitable for capturing the dynamic interplay between formulas and creativity (Myles et al. 1998, 1999) and accounting for the gradual evolution of formulas into

---

<sup>14</sup> Nesselhauf's work is on restricted collocations in written learner language. Consult Author and Colleague (2007) and references cited there for a discussion.

increasingly more productive structures. The role of semi-fixed linguistic templates, i.e., utterance schemas (Tomasello 2000) or formulaic frames (Dabrowska 2000), is crucial in this respect. While semi-fixed expressions of various kinds are recognized as both frequent and important by a wide range of researchers, in L2 studies (e.g. Hakuta 1974; Pawley and Syder 1983; Nattinger and DeCarrico 1992; Lewis 1993; Schmitt and Carter 2000) and corpus linguistics (e.g. Sinclair 1991; Hunston and Francis 2000; Stubbs 2007), it is by way of placing such semi-fixed expressions on a developmental trajectory that UBL is fruitful for exploring the evolution of formulas into more productive patterns. The data for this study, as will become clear in section 3 below, have revealed that semi-fixed templates play a prominent role in ontogenesis.

## **2.2. Can**

Previous longitudinal SLA studies dealing with aspects of *can*-usage focus mostly on request formation (R. Ellis 1992; Achiba 2003) or general communicative competence (Schmidt 1983). A common finding among these studies is that all informants made use of formulas. Schmidt (1983) reported that his informant, Wes, used many formulas, among them *can*-formulas such as *what can I do*, and *can you imagine*, but they were not documented to have an impact on Wes's interlanguage development. Achiba's (2003) case-study on his own 7-year-old daughter learning L2 English seems to show the opposite; she was found to expand on an initial reliance on formulaic patterns to move towards more differentiated request formation; in fact, the acquisition trajectory described by Achiba (2003) looks compatible with the UBL path of item-based learning. Also one of R. Ellis's (1992) informants stands out, as his formulaic use of *Can I have...?* is described as evolving into other *can*-structures over time, suggesting that his learning of this pattern was item-based. These suggestions remain speculative because neither R. Ellis (1992) nor Achiba (2003) investigated development in terms of an item-based learning trajectory. However, both studies seem to suggest that FL is developmentally significant and inextricably coupled with more general linguistic development.

## **2.3. UBL**

Author and Colleague (2007) investigated the development of *do-negation* by a Mexican learner of English and found positive evidence for the UBL path of acquisition. *Do-negation* learning was found to be initially reliant on one specific instantiation of the pattern, *I don't know*. The pattern was gradually expanded to be used with other verbs and pronouns as the underlying knowledge



seemed to become increasingly abstract, as reflected in token and type frequencies. The system emerging in acquisition, then, was initially exemplar-based and development seemed to be based on the gradual abstraction of regularities that link expressions as constructions.

Examining expression of future tense by 16 learners of English, Bardovi-Harlig (2002) made two observations indicating that the UBL path of acquisition is only partially valid for SLA: 1) the role of formulas in initial development was found to be limited; and 2) the use of formulas was found not to diminish ontogenetically. The first observation will be discussed in section 3.1.2. below. The second observation, in fact, does not contradict the suggested path of acquisition insofar as it is connected with the cohabitation in the grammar hypothesis. Most likely Bardovi-Harlig fails to acknowledge this as she investigates the UBL path of acquisition without buying into the whole UBL package. She remains in a framework which compartmentalizes language knowledge, implying that formulaic language is characterized by its display of language knowledge beyond the capabilities of the current interlanguage grammar. Bardovi-Harlig does conclude, however, that N. Ellis's proposed path of acquisition presents a richer view of SLA development than a starting point that excludes formulas.

Mellow (2006) proposed item-based emergentism as an answer to the problem arising from the argument from the poverty of the stimulus for L2 learning. He investigated the acquisition of embedded clauses by Ana, a 12-year old Spanish speaking learner of L2-English, and found that learning in this case was indeed item-based, supporting the UBL trajectory of learning. Ana initially learned the constructions under investigation with particular lexical items and gradually expanded the constructions to apply to new lexical items.

This paper further tests the validity of the UBL trajectory of learning. As it does so via an investigation of the emergence of *can*-patterns in an L2 inventory, it will be discussed to what extent type and token frequencies in learner production can be said to determine or reflect schematicity of learner language representation.

### **3. The present study**

This is a longitudinal case study on Carlos<sup>15</sup>, a Mexican-Spanish speaking classroom learner of

---

<sup>15</sup> Carlos is a pseudonym.

English as a second language (ESL). The data source is the Multimedia Adult English Learner Corpus (MAELC),<sup>16</sup> which consists of audio-visual recordings of classroom interaction. The classrooms, in which the recordings were made, were equipped with video cameras and students were given wireless microphones on a rotational basis; the teacher also wore a microphone (Reder et al. 2003; Reder 2005). The final database of the inquiry consists of transcripts from approx. 40 sessions from September 2001 through February 2005, each consisting of three hours of recordings, in which Carlos is either wearing or sitting next to someone wearing a microphone. Carlos was in his late 20s to early 30s when attending class. During his years in class, Carlos gradually progressed, by the standards of this language program, from Level A to Level D (beginning to intermediate; for more information on the proficiency levels, see Reder 2005).

### 3.1 Data analysis and discussion

Presenting the data while discussing them along the way, this section is organized in a movement from concrete patterns and MWEs via utterance schemas to an investigation and discussion of the abstract linguistic knowledge thought to underlie language use in UBL.

#### 3.1.1 Early patterns and MWEs

An investigation of the UBL trajectory of learning takes MWEs as its natural starting point. The picture for *can*-pattern development does not immediately comply with this *modus operandi*; no dominant MWE can be identified as instigating the learning of a general *can*-pattern. Instead, the data reveal a range of interrelated, locally recurring specific instantiations of a limited number of initially occurring *can*-patterns. Table 1<sup>17</sup>, showing the number of main verb instantiations for each pattern, gives an overview of this early emergence of *can*-patterns. Chronology is represented both horizontally and vertically, with *I can write* as the first and *you can do* as the last instantiation of these patterns to appear in this section of the data. All four patterns in Table 1 are reused at later points in time. At this stage, *I can write* is the only instantiation of *I can verb*, whereas the varied main verb use in *can you Verb*, *can I Verb*, and *you can Verb* suggests that, already at this early

---

<sup>16</sup> MAELC was compiled and is maintained at The National Labsite for Adult ESOL (known locally as the Lab School). The Lab School was supported, in part, by grant R309B6002 from the Institute for Education Science, U.S. Dept. of Education, to the National Center for the Study of Adult Learning and Literacy (NCSALL) and was a partnership between Portland State University and Portland Community College. I am indebted to Steve Reder, John Hellermann, and all the staff at the Lab School without whom this research would not be possible.

<sup>17</sup> The years in the Tables throughout the paper denote the following recording periods: 2001: September – December 2001; 2002: January – May 2002; 2003: September 2003 – February 2004; 2004: September – December 2004; 2005: January – February 2005. Carlos did not attend ESL class consistently, hence the gaps in the recording periods.

stage, these patterns qualify as utterance schemas. Utterance schemas, however, are the focus of a later section; for now the focus is on specific expressions.

	Write	Spell	Show	See	Take	Call	Sit	Say	Ø	Do
<i>I can Verb</i> , 2001	2									
<i>Can you Verb</i> , 2001	2	1	1	1						
<i>Can I Verb</i> , 2002				3	3	2				
<i>You can Verb</i> , 2002							1	3	1	5

Table 1: Emergence of can-patterns

Table 1, showing the totality of Carlos's *can*-usage at this stage in development, indicates that *I can Verb* and *Can you Verb?* are seemingly driven forward in development each by its own specific instantiation, namely *I can write / Can you write?*, respectively. Both may be thought of as MWEs. In 2001, *I can write* is considered a MWE because it is reused by Carlos and because the *I can Verb*-pattern is not used with other main verbs until later. *Can you write* is considered a MWE for the same reasons; because it is the only *can you Verb*-instantiation that recurs, and because it precedes, in the data, the other *can you Verb*-instantiations ontogenetically. Table 1 also displays early emergence of *can I Verb* and *you can Verb* to be centred on a few, mostly recurring instantiations.

The manner in which the patterns emerge, seemingly each other's stepping stone in development, shows how the MWEs are an integral part of the developing linguistic inventory. They seem to represent the backbone of pattern acquisition in the 'construct-i-con', with the patterns forming interrelated, temporary subsystems of sedimentations of experienced language (Hopper 1998). The interrelated systematicity is evident in the overlapping of main verbs and in the stepping-stone metaphor which captures the idea that the emergence of new patterns builds on previously experienced patterns (N. Ellis and Larsen-freeman 2006). Ultimately, the primordial stepping-stone in this development seems to be the MWE *I can write*, the very first *can*-instantiation deployed by Carlos. Therefore, the traditional view of formulaic language, according to which the formulas are uniquely distinguished, insular units beyond the current combinatorial interlanguage system, does

not seem to concur with these data. Rather, at this relatively early stage in acquisition, general pattern development seems to hinge on and be intertwined with a few interrelated MWEs.

It has been suggested that the MWEs are local phenomena. Situated in certain interactional environments, reflecting Carlos's changing communicative needs in the classroom, they seem to be transitory in nature, dependent on recurring usage events. This is substantiated by the following extracts which illustrate the locally situated nature of the MWEs *I can write* (extracts 1a-b), *can you write* (extracts 2a-c), *can I see* (extracts 3a-b), and *you can do* (extract 4). The extracts display how *I can write*, *can you write*, and *can I see* are used in comparable contexts which recur several months apart, whereas *you can do* is only used in one specific interaction. The transitory nature of the MWEs, then, may be seen as a result of Carlos seemingly adapting his linguistic resources (Larsen-Freeman 2006) to the ever-changing challenges of the classroom.

Extract 1a<sup>18</sup>:

(Oct 04 2001)

1 T            okay so check your name very good everybody did okay (.) write the date  
2            for today yesterday and tomorrow (.) Rosalba wake up (.) you are so tired  
3 unknown    xx[x (1) yesterday and tomorrow  
4 T            [okay the date (1) for today  
5 multiple    ((partially in overlap)) today tuesday today tuesday today  
7 C →        ((raises hand)) I can write  
8 unknown    oh today (1) mhm  
8 T            ((walks toward the board, waves pen))  
9 C →        I can write  
10 T          you always write Carlos ((gives Carlos the pen))

Extract 1b

(May 17 02)

---

<sup>18</sup> All data extracts are accessible online: <http://www.labschool.pdx.edu/Viewer/viewer.php?pl=sweCaLcan>. Clip length may exceed transcripts.

Transcript conventions: C = Carlos; T = Teacher; P(1,2) = Partner(1,2); xxx = inaudible; [] = overlap; (1/2)/(.) = 1/2 second(s) pause / very short pause; ((yyy)) = transcriber's comments. Arrows mark the target expressions.

1 T           okay (.) now uh I need two students (.) here's one  
 2            (.)  
 3 T           where's the other one  
 4 C →       ((gets up)) I don't know (.) but I can write (.) write ((reaches for Teacher's pen))  
 5 T           ((gives pen to another student))  
 6 C           oh he? ((points to student, goes back to desk))

Extracts 12-b show that *I can write* is employed in contexts where Carlos volunteers to write on the board. In extract 1a Carlos is selected by the teacher, whereas in extract 1b the teacher seemingly selects two other students. In both cases, Carlos makes a physical move (raises hand / gets up) and volunteers, expressed by *I can write*. This MWE only becomes relevant in these contexts, which recur 7 months apart, suggesting that the MWE and the usage event are somehow intertwined.

Extract 2a:

(Oct 15 01) (Situation: a new student from Laos has just come into the class. She sits down behind Carlos and his Vietnamese partner who are preoccupied with writing in Vietnamese. The teacher gives the new student a piece of paper to be used as a name tag and asks Carlos to tell her what to do next.)

1 T           so you're learning Vietnamese instead of English (.) m'kay tell her what to do with  
               this ((pointing to name tag)) What does she need to do?  
 2            (1)  
 3 C           Eh  
 4 T           [tell her  
 5 C           [what  
 6 T           ((picks up C's name tag)) tell her  
 7 C           alright ((starts turning to new student))  
 8 T           write your first name in big letters  
 9 P           small  
 10 T          mhm  
 11 C          I'll say she say she

- 12 T            please (.) uhuh tell her uhuh  
 13            (.)  
 14 C →        ahhah can you write here eh (.) your first name?  
 15            (.)  
 16 C            what is your first name?

The interaction proceeds for some time until the students and the teacher have jointly managed to get the new student's name on the name tag, and the action is accomplished.

Extract 2b:

(Oct 15 01) (Ongoing task: Students are asked to talk to three peers and ask them their names and where they come from. Note that *can* is not among the forms to be practiced.)

- 1 C            what is your first name?  
 2            (.)  
 3 P            it's chandra (.) chandra seh  
 4 C            mm?  
 5 P            chandra  
 6 C            chandra?  
 7 P            see  
 8 C            jann  
 9            (.)  
 10 P          xx[x  
 11 C →        [nn ah I don't know xxx ah no help yiaa (2) can you write here (2) your name? (3)  
               ooh ((fiddling with papers, writes)) chans ((writes)) ah your last last name?

Extract 2c:

(Apr 19 02)

- 1 C →        eh can you write the the (.) her name in in you ((motioning writing))  
 2            (1)  
 3 P1          chinese?

4 C uh ye yeah  
5 P1 Marie ((pointing to P2))  
6 C uhuh  
7 P1 ((writes)) marie marie  
8 C ((points to what P1 wrote)) hm  
9 P2 ((looks at what P1 wrote, smiling))  
10 C xxx  
11 P2 wow <spn>

Extracts 2a-c show how in three comparable situations, two of them happening on the same day and the third one six months later, Carlos uses *can you write* for a specific purpose, namely as a request for partners to write down names. Depending on the on-going interactions he may have different reasons for doing so (asked by teacher (2a), prompted by task (2b), or out of curiosity (2c)<sup>19</sup>), but the employment of the MWE as a request for a peer to write down a name remains stable throughout, further substantiating the suggestion that the interactional context and the relevance of the MWE are interdependent.

Extract 3a:

(Jan 29 02)

1 C xxx picture  
2 P ((hands C her pictures)) uh sho[w hh]  
3 [ehn?]  
4 P eh show heh[hehheh]  
5 C [hehheh]  
6 P eh show his picture (.) hn  
7 (2)  
8 C ((fiddling with his things)) what picture  
9 (1)  
10 P I sh yeah show me [the  
11 C [oh xxx must be here  
12 (5)

---

<sup>19</sup> This is speculation, but Carlos seemingly has an affection for Asian writing systems, cf. situation prior to extract 2a.

- 13 P e[h  
 14 [ah my friends  
 15 P oh heh[hehhehheh]  
 16 C [hmhmhmhm]]  
 17 P ((looking at C's pictures)) mhm mhm (2) o[i (.) hm ]  
 18 C → [hehhehheh] hey lemme see (.) ai can I see?  
 ((taking P's pictures))  
 19 P ((nods))  
 20 C yeah?  
 21 P ((nods)) xxx mhm

Extract 3b:

(Jun 06 02) (situation: End of term. P and C talk about what class to attend the following term. Some students need to fill in application forms for a variety of reasons; C is curious to know more about P's forms. She has two such forms.)

- 1 C → why do you have two? (.) can I see?  
 2 P no  
 3 C why not?  
 4 P ((begins to put forms away))  
 5 C xxx  
 6 P xxx ((puts forms away))

Extracts 3a and 3b show how Carlos uses the MWE *can I see*<sup>20</sup> as a specific request to see something in the co-participant's possession. All extracts above show how the MWEs, their respective functions seemingly stable over time, are locally contextualized, i.e., they become relevant in certain recurring local interactional contexts in the classroom; specific interactional

---

<sup>20</sup> There is a third comparable situation in which the MWE is employed, on Jan 18 02. It is not transcribed here because the only completely audible turn is Carlos using the MWE to ask to see something in his partner's book. His partner then complies and helps him find what he wants to see. In other words, this situation matches the ones in the extracts. Please note that the other MWE for the *can I*-pattern, *can I take*, is excluded from interactional analysis. Carlos, when asked by the teacher to collect student journals, uses the MWE to request permission to take their journals. While the students display understanding of the request inasmuch as they comply with it by handing over the journals, Carlos receives no verbal answer on any of the three occasions of use. *Can I call*, the final expression in this pattern, is used twice in the interaction as Carlos requests a peer for permission to call her. The extract is left out in the interest of saving space.



activities and employment of specific MWEs co-inform each other. In UBL terms, certain usage events seem to be coupled with certain linguistic expressions.

This correlation between communicative situation and linguistic expression is not a new discovery, however. Nattinger and DeCarrico (1992), Weinert (1995), and Bardovi-Harlig (2002) mention situational dependence or functional purpose as a defining criterion for FL. Wray (2002: 262)<sup>21</sup> similarly holds that functionality is 'the most fundamental aspect of formulaicity', and Kecskes's (2003) situation-bound utterances are directly based on the idea that certain formulas are used for specific purposes. Also L2 socialization research has established a correlation between linguistic and social routines (e.g. Kanagy 1999; Hall 2004).

Viewed in this light, the emergent nature of the patterns revealed in Table 1 is seen as a reflection of the mutually dependent nature of interactional requirements and relevance of specific MWEs. This indicates that emergentism is not only a matter of incipient patterns developing in a stepping-stone fashion, but also a matter of shifting and adapting behaviours in distributed environments calling upon linguistic patterns which might be stable or in flux (Thelen and Bates 2003; Larsen-Freeman 2006). The MWEs investigated here seem mostly to be in flux, invoked, as it were, by certain interactional circumstances, showing 'the mutual constituency of linguistic resources and tasks' (Firth and Wagner 2007: 812). Interactional activity and emergent linguistic knowledge therefore need to be co-investigated, and use and acquisition viewed as essentially inseparable (Firth and Wagner 1998; Larsen-Freeman 2006), in longitudinal studies on interactional L2 data.

Keeping in mind that Table 1 showed both initial occurrence of MWEs and pattern development in a stepping-stone fashion, the locally contextualized MWEs may be thought of as the backbone of the increasing productivity of the patterns in question. Their presence guides the on-going development of the linguistic inventory, as Carlos constructs his L2. This importance of MWEs in ontogenesis, however, only becomes apparent if locally anchored uses are investigated. Bardovi-Harlig (2002: 192), stating that initial use of formulas may be 'so brief that it cannot be detected in the corpus', may have missed the importance of the initial formulas precisely because she does not distil local uses from the totality of her data.

---

<sup>21</sup> I thank an anonymous reviewer for directing my attention to this quote in Wray (2002).



- 27 P            yeah
- 28 C→        why not you can do
- 29 P            becau becau um this morning nu:h I did I did nu:h I did do (.)
- 30            hahhah[hahhah. ]
- 31 C            [hahhahhah] okay

### 3.1.2 Types and tokens.

So far, the analysis has dealt with tokens; i.e., specific instantiations of the *can*-patterns. This section moves up the ladder of schematicity to investigate type and token frequencies and their impact on *can*-pattern development. Table 2 below shows type and token frequencies of the four *can*-patterns that were investigated in the previous section. Tokens are the total number of instantiations, types are the number of different instantiations, and the type-token ratio gives the relationship between the two. The more tokens, the more frequently a pattern is used. The more types, the more different instantiations of the patterns are used. A productive pattern, then, yields a high type-token ratio. Table 2 therefore shows, as already suggested, that *I can Verb* has not evolved into an utterance schema yet, whereas the other three patterns display utterance schema-like behaviour, as it were.

	Tokens	Types	Ratio
<i>I can Verb</i> , 2001	2	1	--
<i>Can you Verb</i> , 2001	5	4	0,80
<i>Can I Verb</i> , 2002	8	3	0,38
<i>You can Verb</i> , 2002	10	4	0,40

Table 2: Token and type frequencies of initially recurring *can*-patterns.

The high ratio for *can you Verb* reflects the existence of only one MWE, *can you write*, and the relatively low ratios for *can I verb* and *you can verb* reflect their initial dependence on several MWEs, as explored in the previous section. This section explores the productivity of the patterns, not as specific expressions but as utterance schemas. Whereas the MWE was operationalized as a recurring string of words used for a coherent communicative purpose, the utterance schema is a partially schematized pattern with an open slot, not unlike Pawley and Syder's (1983) lexicalized

sentence stems and Sinclair's (1991) semi-pre-constructed phrases.

	2001	2002	2003	2004	2005
<u>I can Verb</u>	MWE	5/5	4/5	4/4	3/3
<i>Can you Verb</i>	4/5	3/3	--	--	--
<i>Can I Verb</i>	--	3/8	2/2	1/1	--
<i>You can Verb</i>	--	3/9	1/1	1/1	--
<i>You can't Verb</i>	--	--	1/1	1/2	--
<i>They can't Verb</i>	--	--	1/1	--	--
<i>I can't Verb</i>	--	--	1/1	2/2	--
<i>We can Verb</i>	--	--	--	1/1	1/1

Table 3: Emergence of utterance schemas; numbers show type and token frequencies.

Table 3, displaying an overview of Carlos's *can*-usage development in toto, shows initial development to be characterized by one MWE *I can write* and one utterance schema *can you Verb*. Subsequent development sees a general expansion in utterance schema production. L2 learning as the process of constructing another language may, in this case, be described as the gradual expansion of varied utterance schema use. The development of utterance schemas, as depicted in Table 3, displays the same stepping-stone tendency as evident in Table 1. Additions to the linguistic inventory are always traceable to previously experienced language. The four initially occurring *can*-patterns, which the previous section showed to have emerged on the basis of interrelated MWEs, seem to constitute the foundation from which the other utterance schemas emerge.

The data also demonstrate that development is not entirely linear. The four original patterns vastly dominate the picture, some patterns are highly portable (*I can verb, you can Verb, can I verb*), and some dramatically diminish (*you can Verb, can I verb, can you verb*). In other words, all linguistic patterns appear and disappear in the data; only the time-span in terms of which this fluctuating nature of language use and knowledge is apparent differ. MWEs, coupled with certain interactional circumstances, seem to be more temporally fragile than the utterance schemas which, because of their partial schematicity, are less susceptible to environmental changes. This is especially evident in terms of the *I can Verb*-schema which is in use throughout development. However, the non-linearity described here is different from the processes of forgetting and backsliding, i.e., regressing

to earlier stages, which characterise U-shaped development (e.g., McLaughlin 1990; Larsen-Freeman 1991). Rather, this non-linearity resembles the waxing and waning of linguistic patterns in an iterative process of learning where learners recycle known material in a spiralling process (Ellis and Larsen-Freeman 2006; Larsen-Freeman 2006; de Bot et al. 2007).

To sum up, the data have shown development to be traceable in terms of emergentism which holds that 'utterances are closely similar to previous utterances and that anything that is said has been said in something like that form before' (Hopper 1998: 165). As Carlos constructs his L2 he seems to build on previously encountered material. His linguistic experience consists of sediments of previous interactions, which renders L2 learning an iterative process of revisiting known territories (Larsen-Freeman 2006). Early patterns were thus shown to be traceable to transitory MWEs, and later patterns to earlier patterns, displaying a constant development toward a richer and more varied linguistic inventory in response to richer and more varied interactional needs (Lantolf and Thorne 2006). Language learning in this sense of emergentism is a matter of 'expanding a repertoire of communicative contexts' (Hopper 1988: 171).

### 3.1.3 Overall *can*-usage frequencies.

Previous sections have dealt with interrelated but separate MWEs and utterance schemas. This section investigates the possible existence of a completely abstract schematized *can*-construction in my focal student.

	Tokens	Types	Ratio
2001	7	5	0,71
2002	25	14	0,56
2003	11	10	0,91
2004	11	10	0,91
2005	4	4	1,00

Table 4: Type and token frequencies for *can*-usage.

Table 4 displays the totality of Carlos's *can*-usage. No distinction is made among utterance schemas; types are differentiated on the basis of main verb use. The ratios in Table 4 show a tendency towards an increase in productivity from 2002 onwards. This increase in productivity is

partly a result of (perhaps pattern-specific) lexical learning; i.e., main verb production in *can*-patterns becomes increasingly varied. Referring back to Table 3, however, it is evident that another source of the very high ratios in 2003 onwards is Carlos's use of different utterance schemas. Carlos, then, is working on a very useful linguistic inventory of slightly increasing productivity, but an inventory which in terms of underlying abstract representation does not seem to undergo fundamental changes. Rather, increased productivity is a result of utterance schema additions to the linguistic inventory.

The fairly high type-token ratio throughout development reflects a high degree of productivity based on utterance schemas rather than movement towards a fully abstract construction. Thus, 'type' is not necessarily the frequency equivalent to a fully schematic feature of language knowledge; instead, it may need to be considered a parallel to an utterance schema. The data therefore suggest that, in order for the hypothesis on frequency and acquisition to hold true (token frequency = storage as concrete whole; type frequency = storage as abstract, schematic construction), it is necessary to adjust the relationship between type frequencies and pattern use and acquisition to accommodate utterance schemas; i.e., instantiations that are not, seen from the perspective of adult native linguistic competence as envisioned in UBL, fully schematic.

### 3.1.4 Abstractness and linguistic knowledge in SLA

As already suggested, describing linguistic development in terms of increasing underlying abstractness may be a dubious enterprise. Putting this observation to the test in terms of the present data, this section investigates such development along two lines: 1) number of different pronouns used in subject position; and 2) tense variation. Showing pronoun and tense use across time in the *can*-pattern types declarative, interrogative, and negative, Table 5 confirms what was also apparent in Table 3, namely that certain pronouns at certain points in time seem to be more entrenched than others, with 1<sup>st</sup> and 2<sup>nd</sup> person singular generally dominant from 2002 through 2004. However, in terms of pronoun development, no clear expansion towards increased productivity is detectable, with different pronouns in use in different patterns at various times.

	2001	2002	2003	2004	2005
1 <sup>st</sup> pers., sing.	Decl.	Decl. Interrog.	Decl. Interrog.	Decl. Interrog.	Decl.

			Neg.	Neg.	
2 <sup>nd</sup> pers., sing.	Interrog.	Decl. Interrog.	Decl. Neg.	Decl. Neg.	
3 <sup>rd</sup> pers., sing.					
1 <sup>st</sup> pers., pl.				Decl.	Decl.
2 <sup>nd</sup> pers., pl.	Interrog.				
3 <sup>rd</sup> pers., pl.			Neg.		
Past tense					

Table 5: Abstractness development?

With pronoun use and pattern functions thus distributed, and given the lack of tense variation in the pattern, it is empirically sound and therefore plausible to assume that the linguistic inventory as presented here consists of a range of interrelated utterance schemas, revolving around the concrete employment of *can* itself. Carlos's linguistic inventory becomes increasingly productive and internally coherent as the patterns emerge ontogenetically, with later additions to the inventory always traceable to previous linguistic experience. Carlos's language learning as investigated here does not seem to be measurable along the lines of increasing syntactic combinability, and, acknowledging the distributed nature of the *can*-pattern instantiations, the *PRN*-slot cannot be empirically substantiated to be sanctioned by a fully schematic representation. The data, in other words, do not support the idea of an ultimately schematic underlying representation, such as a *NP Can*-construction or a *PRN Can*-construction, sanctioning all *can*-instantiations.

Such ultimately schematic representations might be the psychological endpoint for adult L1 users, or they might not, but this final level of abstract constructional knowledge remains unsupported by these L2 data. The completion implied by the final, fully schematic level in the proposed item-based path of language acquisition in UBL therefore seems misguided for the purposes of longitudinal studies of SLA. This does not necessarily imply that L2 learners never reach this level of constructional mastery as predicted by the UBL path of learning; intriguingly, in Author and Colleague (2007) we did find evidence for fully abstract negation knowledge in our study on the same focal student. It does, however, raise the question of what constitutes an 'endpoint' on the item-based path of learning, and of how abstract underlying linguistic knowledge ultimately is. It

seems that, in contrast to the *do-neg*-pattern, the focal construction in this paper does not easily lend itself to abstraction. Therefore, it becomes an empirical question to determine the degree of underlying abstractness for each construction under investigation. The present study suggests that it is safer to start a longitudinal exploration from the most concrete starting point possible; the degree of abstractness development, and thus a possible structural endpoint of learning, must be an empirical issue. The present data show no evidence of a journey along the interlanguage continuum toward structural completion and increasingly native-like conformity. Extract 4 above also supports this, showing non-native-like and native-like uses of the same expression (*you can do, you can do it, why not you can do*) within 1 minute and 15 seconds of interaction. Rather, the ontogenetic process of constructing another language seems to be a constant process without a visible endpoint (Firth and Wagner 1998; Lantolf and Thorne 2006), dominated by 'the sedimentation of frequently used expressions and parts of expressions' (Hopper 1998: 166).

While a structural endpoint of learning may be elusive, the starting point is investigable as a matter of exemplar-based learning; all *can*-usage in the data seems to ultimately spring from one particular source, namely the locally situated MWE *I can write*. While there is no guarantee that this was in fact Carlos's primordial use of *can*, the data have shown that it is possible at least to attempt a demonstration of where a given linguistic usage comes from – the litmus test of an emergentist account of development (MacWhinney 2006).

UBL, then, with its proposed item-based path of language learning seems promising, in N. Ellis's words, as a default in guiding investigations into longitudinal L2 learning. However, its operationalization of performance as a subset of competence (Barlow and Kemmer 2000) is a problematic *modus operandi* for SLA. Use in SLA cannot be thought of as a subset of knowledge because we do not know exactly what might constitute an exhaustive account, at the most abstract level of schematicity, of the linguistic resources of a given L2 learner. In other words, the most abstract atemporal categories most often preferred by theoretical linguistics, proponents of UBL or otherwise, are insufficient to account for the emergent L2. A corner-stone in the notion of 'item-based learning', this observation is parallel to Tomasello's (2000, 2003) point that the incipient linguistic inventory of children learning their L1 should not be described and analysed with recourse to the categories of an adult mature native speaker because young children and adults do not operate on identical linguistic units (see also Peters 1983). Thus, questioning the validity of



applying an essentially timeless and rigorous theoretical framework to something that is fundamentally dynamic and distributed (Larsen-Freeman 2006), this paper advocates the future study of L2 linguistic development in terms of an empirically grounded locally contextualized grammar, consisting of flexibly abstract units of actual language use in interaction.

#### **4. Conclusions and implications**

The proposed item-based path for learning an L2 seems to be valid, albeit with one major adjustment. Formulas, or MWEs, should be seen as interactionally and locally contextualized. This means that they may possibly be transitory in nature; i.e. their deployment over time is occasioned by specific usage events. Such events must be recurrent in order for the MWEs, at least the ones identified here, to be retained by the learner over time. A further observation has been made which does not require an adjustment but is already accommodated in UBL and emergentism, namely that productivity enhancement is partially concrete, based on utterance schema development, and traceable to previous experience.

It follows from both these observations that ontogenetic language development is inextricably coupled with language use. A full theory or model of SLA, then, must incorporate room for studying these local contexts in a more detailed manner to investigate in depth the interplay between local interactional contingencies and portable linguistic experience. One question for future research lies in delineating the characteristics of those aspects of language knowledge that are situated and transitory and those that are durable and portable. This research has indicated that MWEs may be generally transitory and locally contextualized whereas more schematic language knowledge may be less susceptible to environmental changes and thus more sturdy in its portability.

In terms of the hypothesis on frequency and acquisition (token frequency = MWE acquisition; type frequency = more abstract ('constructional') acquisition), it seems necessary for SLA purposes to adjust the relationship between type frequencies and pattern use and acquisition to accommodate utterance schemas; i.e., instantiations that are not fully schematic. Such utterance schemas rather than fully abstract constructions were shown here to guarantee productivity, which led to a more principled discussion of what constitutes a schematic endpoint of L2 learning.

While questions of endpoints of learning have not been ultimately answered in this paper, the

starting point seems to be a clearer matter. The data have shown the emergence of *can*-usage in terms of the initial importance of MWEs acquisitionally, their local importance communicatively, and their reflexive dependence on usage events. MWEs were shown not to be reused by Carlos over a longer time-span, and accordingly questions of portability were not seen as definable a priori. The learning of the items that did seem to be definable in terms of portability, namely utterance schemas, was not describable along the lines of increasingly complex structures along a route from non-native-like to (near)native-like structural mastery, nor was it ultimately a matter of increased abstractness in terms of underlying psychological representations. While these two dimensions also have a part to play in determining language learning longitudinally from a UBL perspective, as shown in Author and Colleague (2007), the investigative road of analysing learner language must start from the concrete experience of the learner and take as its point of departure the idea that the learning of L2 structure may not have a visible endpoint. Focusing on the 'usage' of Usage-Based Linguistics, I further encourage future research to investigate in detail the link between local interactional phenomena and linguistic development. Otherwise, we will lose the richer picture of the incremental nature of the emergent linguistic inventory as learners go through the process of constructing another language.

## References.

**Achard, M.** 2007. 'Usage-based semantics: Meaning and distribution of three French 'breaking' verbs' in Nenonen and Niemi (eds.)

**Achiba, M.** 2003. *Learning to Request in a Second Language: A Study of Child Interlanguage Pragmatics*. Clevedon: Multilingual Matters.

**Author and Colleague** 2007. 'Are recurring multi-word expressions really syntactic freezes? Second language acquisition from the perspective of usage-based linguistics' in Nenonen and Niemi (eds.). URL:

<http://www.labschool.pdx.edu/research/current/Eskildsen%20&%20Cadierno,%202007.pdf>

**Bardovi-Harlig, K.** 2002. 'A new starting point? Investigating formulaic use and input in future expression.' *Studies in Second Language Acquisition* 24/2: 189-198.

**Bardovi-Harlig, K.** 2006. 'On the role of formulas in the acquisition of L2 pragmatics' in Bardovi-Harlig, K., C. Félix-Brasdefer, & A. Omar (eds.): *Pragmatics and Language Learning, Vol. 11* Honolulu, HI: National Foreign Language Resource Center, University of Hawai'i at Manoa.

**Barlow, M. and S. Kemmer.** 2000. 'Introduction: A usage-based conception of language' in Barlow and Kemmer (eds.).

- Barlow, M. and S. Kemmer.** 2000. *Usage-Based Models of Language*. Stanford: CSLI.
- Bohn, O.-S.** 1986. 'Formulas, frame structures, and stereotypes in early syntactic development: Some new evidence from L2 acquisition.' *Linguistics* 24: 185-202.
- Bolander, M.** 1989. 'Prefabs, patterns and rules in interaction' in Hyltenstam, K. and L. K. Obler (eds.): *Bilingualism Across the Life Span*. Cambridge: Cambridge University Press.
- Bolinger, D.** 1979. 'Meaning and memory' in Haydu, G. G. (ed.): *Experience Forms. Their Cultural and Individual Place and Function*. The Hague: Mouton.
- Butler, C. S.** 2003. 'Multi-word sequences and their relevance for recent models of Functional Grammar.' *Functions of Language* 10/2: 179–208.
- Dabrowska, E.** 2000. 'From formula to schema: The acquisition of English questions.' *Cognitive Linguistics* 11/1-2: 83-102.
- Dabrowska, E. and E. Lieven.** 2005. 'Towards a lexically specific grammar of children's question constructions.' *Cognitive Linguistics* 16/3: 437-474.
- De Bot, K., W. Lowie and M. Verspoor.** 2007. 'A Dynamic Systems Theory approach to second language acquisition.' *Bilingualism: Language and Cognition* 10/1: 7-21.
- Ellis, N. C.** 2002. 'Frequency effects in language processing – a review with implications for theories of implicit and explicit language acquisition.' *Studies in Second Language Acquisition* 24/2: 143-188.
- Ellis, N. C. and D. Larsen-Freeman.** 2006. 'Language emergence: Implications for applied linguistics – Introduction to the Special Issue.' *Applied Linguistics* 27/4: 558-589.
- Ellis, R.** 1992. 'Learning to communicate in the classroom: A study of two learners' requests.' *Studies in Second Language Acquisition* 14: 1-23.
- Firth, A. and J. Wagner.** 1998. 'SLA Property: No Trespassing!' *The Modern Language Journal* 82/1: 91-94.
- Firth, A. and J. Wagner.** 2007. 'S/FL learning as a social accomplishment: Elaborations on a 'reconceptualised' SLA.' *The Modern Language Journal* 91, 800-819.
- Goldberg, A.** 2003. 'Constructions: A new theoretical approach to language.' *Trends in Cognitive Sciences* 7/5: 219-224.
- Granger, S.** 1998. 'Prefabricated patterns in advanced EFL writing: collocations and formulae' in A. P. Cowie (ed.): *Phraseology. Theory, analysis, and applications*. Oxford: Clarendon Press.
- Hakuta, K.** 1974. 'Prefabricated patterns and the emergence of structure in second language acquisition.' *Language Learning* 24: 287-297.

- Hall, J. K.** 2004. 'Practicing Speaking in Spanish: Lessons from a High School Foreign Language Classroom' in Boxer, D. (ed.): *Studying Speaking to Inform Second Language Learning*. Clevedon: Multilingual Matters.
- Hoey, M.** 2007. 'Grammatical creativity: a corpus perspective' in Hoey et al. (eds.)
- Hoey, M., M. Mahlberg, M. Stubbs and W. Teubert.** 2007. *Text, Discourse and Corpora*. London: Continuum.
- Hopper, P. J.** 1988. 'Emergent grammar and the a priori grammar postulate' in Tannen, D. (ed.): *Linguistics in Context*. Norwood: Ablex.
- Hopper, P. J.** 1998. 'Emergent grammar' in Tomasello, M. (ed.): *The New Psychology of Language, volume 1*. Mahwah, NJ.: Lawrence Erlbaum.
- Hunston, S. and G. Francis.** 2000. *Pattern Grammar. A Corpus Driven Approach to the Lexical Grammar of English*. Amsterdam: John Benjamins.
- Irujo, S.** 1993. 'Steering clear: Avoidance in the production of idioms.' *International Review of Applied Linguistics in Language Teaching* 31/3: 205-219.
- Kanagy, R.** 1999. 'Interactional routines as a mechanism for L2 acquisition and socialization in an immersion context.' *Journal of Pragmatics* 31: 1467-1492
- Kecskes, I.** 2003. *Situation-Bound Utterances in L1 and L2*. Berlin: Mouton de Gruyter
- Krashen, S. D. and R. Scarcella.** 1978. 'On routines and patterns in language acquisition and performance.' *Language Learning* 28: 283-300
- Langacker, R. W.** 1987. *Foundations of Cognitive Grammar, vol. 1. Theoretical prerequisites*. Stanford: Stanford University Press.
- Langacker, R. W.** 2000. 'A Dynamic Usage-Based Model' in Barlow and Kemmer (eds.)
- Lantolf, J. P. and S. L. Thorne.** 2006. *Sociocultural Theory and the Genesis of Second Language Development*. Oxford: Oxford University Press.
- Larsen-Freeman, D.** 1991. 'Second Language Acquisition research: Staking out the territory.' *TESOL Quarterly* 25/2: 315-350.
- Larsen-Freeman, D.** 2006. 'The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English.' *Applied Linguistics* 27/4: 590-619.
- Lewis, M.** 1993. *The Lexical Approach*. UK: Language Teaching Publications
- MacWhinney, B.** 2004. 'A multiple processes solution to the logical problem of language acquisition.' *Journal of Child Language* 31: 883-914

- MacWhinney, B.** 2006. 'Emergentism – use often and with care.' *Applied Linguistics* 27/4: 729-740.
- McLaughlin, B.** 1990. 'Restructuring.' *Applied Linguistics* 11/2: 113-128
- Mellow, J. D.** 2006. 'The emergence of second language syntax: A case study of the acquisition of relative clauses.' *Applied Linguistics* 27/4: 645-670
- Myles, F., J. Hooper, and R Mitchell.** 1998. 'Rote or rule? Exploring the role of formulaic language in classroom foreign language learning.' *Language Learning* 48/3: 323-363
- Myles, F., R. Mitchell, and J. Hooper.** 1999. 'Interrogative chunks in French L2. A basis for creative construction?' *Studies in Second Language Acquisition* 21: 49-80
- Nattinger, J. P. and J. S. DeCarrico.** 1992. *Lexical Phrases and Language Teaching*. Oxford: Oxford University Press
- Nenonen, M and S. Niemi (eds.).** 2007. *Collocations and Idioms 1: Papers from the First Nordic Conference on Syntactic Freezes, Joensuu, May 19-20, 2007*. Studies in Languages, University of Joensuu, Vol. 41. Joensuu: Joensuu University Press
- Nesselhauf, N.** 2005. *Collocations in a Learner Corpus*. Amsterdam: John Benjamins.
- Pawley, A. and F. Syder.** 1983. 'Two puzzles for linguistic theory: Nativelike selection and nativelike fluency' in Richards, J. C. and R. W. Schmidt (eds.): *Language and communication* London: Longman
- Peters, A. M.** 1983. *Units of Language Acquisition*. Cambridge: Cambridge University Press
- Read, J. and P. Nation.** 2004. 'Measurement of formulaic sequences' in Schmitt (ed.).
- Reder, S.** 2005. 'The "Lab School".' *Focus on Basics* 8a. URL: [http://www.ncsall.net/fileadmin/resources/fob/2005/fob\\_8a.pdf](http://www.ncsall.net/fileadmin/resources/fob/2005/fob_8a.pdf)
- Reder, S., K. A. Harris, and K. Setzler.** 2003. 'A Multimedia Adult Learner Corpus,' *TESOL Quarterly* 37/3: 546-557.
- Schmidt, R.** 1983. 'Interaction, acculturation, and the acquisition of communicative competence: A case study of an adult' in Wolfson, N. and E. Judd (eds.): *Sociolinguistics and Language Acquisition*. Rowley, Mass.: Newbury House
- Schmitt, N.** 2004. *Formulaic Sequences*. Amsterdam: John Benjamins.
- Schmitt, N. and R. Carter.** 2000. 'Lexical phrases in language learning.' *The Language Teacher*. URL: <http://www.jalt-publications.org/tlt/articles/2000/08/schmitt>
- Schmitt, N. and R. Carter.** 2004. 'Formulaic sequences in action: an introduction' in Schmitt (ed.).

- Sinclair, J.** 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press
- Stefanowitsch, A. and S. Th. Gries.** 2003. 'Collostructions: Investigating the interaction of words and constructions.' *International Journal of Corpus Linguistics* 8/2: 209-243.
- Stubbs, M.** 2007. 'Quantitative data on multi-word sequences in English: the case of the word *world*' in Hoey et al. (eds.)
- Thelen, E. and E. Bates.** 2003. 'Connectionism and dynamic systems: are they really different?' *Developmental Science* 6/4: 378–391.
- Tomasello, M.** 2000. 'First Steps toward a Usage-Based Theory of Language Acquisition' *Cognitive Linguistics* 11/1-2: 61-82.
- Tomasello, M.** 2003. *Constructing a Language*. Cambridge: Cambridge University Press.
- Tummers, J, K. Heylen, and D. Geeraerts.** 2005. 'Usage-based approaches in cognitive linguistics: A technical state of the art.' *Corpus Linguistics and Linguistic Theory* 1-2: 225-261.
- Van Lancker-Sidtis, D. and G. Rallon.** 2004. 'Tracking the incidence of formulaic expressions in everyday speech: Methods for classification and verification. *Language and Communication* 24: 207-240.
- Weinert, R.** 1995. 'The role of formulaic language in second language acquisition: A review.' *Applied Linguistics* 16/2:180-205.
- Wood, D.** 2002. 'Formulaic language in acquisition and production: implications for teaching.' *TESL Canada Journal* 20/1: 1-15
- Wray, A.** 2002. *Formulaic Language and the Lexicon*. New York, NY: Cambridge University Press.
- Wray, A. and M. R. Perkins.** 2000. 'The functions of formulaic language: an integrated model.' *Language and Communication* 20: 1-28

### Chapter 3: ESK2

What's New? - Routines and Creativity along a Usage-Based Path of Second Language Learning<sup>22</sup>

#### Abstract.

Exploring the relationship between routines and creativity in linguistic development, this paper investigates second language (L2) learning in terms of an iterative process in which learners reuse linguistic matter to a large extent (Larsen-Freeman 2006; Author in press). The object of research in longitudinal L2 studies is therefore understood as a hybrid between locally applied usage patterns and application of the same and related usage patterns over time. Investigating the linguistic repertoires of two classroom learners of English, I adopt from Usage-Based Linguistics an item-based path of language learning, from formulas through low-scope patterns to abstract constructions. This makes for an empirically grounded, emergent grammar, consisting of units of spontaneously occurring language use. The results call for a reconceptualisation of the anatomy of L2 inventories, including emergent creativity, as building on recycled linguistic matter, most prominently in the form of "formulas" and "formulaic frames", here referred to as "multi-word expressions" and "utterance schemas", respectively.

#### Introduction: The linguistic inventory and second language development.

Building on insights from Author & Colleague (2007) and Author (in press) this paper argues in favour of viewing second language (L2) development as fundamentally usage-based and item-based. This implies, as is generally done in Usage-Based Linguistics (UBL), abandoning a number of prevalent dualisms in linguistics and applied linguistics, especially those between competence and performance, and lexis and syntax. UBL characterizes language knowledge as a structured inventory of symbolic units, i.e., form-meaning pairings (Langacker 1987). This inventory is thought of as a diverse assortment of linguistic patterns at various levels of complexity and abstractness, allowing for the co-habitation in the grammar of abstract schematized representations and their concrete instantiations (Langacker 2000; Achard 2007).

L1 acquisition in UBL is viewed as fundamentally item-based and has been empirically substantiated to proceed along a trajectory from formulas via low-scope patterns to fully abstract

---

<sup>22</sup> This article is under review for publication in *International Review of Applied Linguistics*.

constructions (Dabrowska 2000; Tomasello 2003; Dabrowska and Lieven 2005). While these researchers fundamentally adhere to the general outline of this proposed path of learning, they differ somewhat on the terminology. Therefore, it is worth pausing to take a closer look at the terms involved in viewing language learning as item-based along this proposed path of development.

Notoriously difficult to capture terminologically, formulas are most often defined as rote-learned chunks (e.g., *Lemme-see*, *I wanna do it*, *I dunno*). The term used here, based on Author and Colleague (2007) and Author (in press), is recurring multi-word expression (MWE), which denotes the usage-based vantage point on at least two accounts: 1) only recurring words strings used for coherent communicative purposes and demonstrably found in the data are included as such – i.e., no patterns are defined as MWEs a priori – and 2) it underlines the abolition of the strict lexis-syntax dichotomy, and, following the cohabitation in the grammar hypothesis (Langacker 2000; Achard 2007), it does not presuppose the idea of excising concrete items from the linguistic inventory as more abstract linguistic representations may be acquired over time.

Low-scope patterns (e.g., *I wanna Verb*, *I don't Verb*) consist of a fixed part and an open slot. In this paper, the term utterance schema, borrowed from Tomasello (2000), will be used to refer to patterns that are partially fixed and partially productive, because some schematic knowledge is seen to sanction the use of given lexical items in the open slot.

The posited level of ultimate abstractness consists of schematic knowledge of symbolic units; e.g., *I don't Verb* is specified more generally in the inventory as the negation construction *NP aux neg Verb* which then sanctions a wide range of lexical specifications. As pointed out by Lieven et al. (2003), however, the question of underlying linguistic representation and thus the abstractness of linguistic knowledge is not easily answered. One kind of evidence to go by is frequency-based; a high type frequency implies that a pattern has been abstracted from concrete pattern usage, a high token frequency implies concrete pattern entrenchment; i.e., linguistic items may be routinized through frequent use (Langacker 2000; Tomasello 2003). The transition from concrete to abstract linguistic representation is of course fluid, and sometimes it is necessary to go beyond frequency effects in language development to explore whether fully abstract schematicity seems to apply to the linguistic inventory under investigation (Author in press). However, at its most abstract and elegant, the system that emerges in development, when this usage-based path of learning is applied



in full to second language acquisition (SLA) research, is characterized as the gradual abstraction of regularities that link expressions as constructions (Author and Colleague 2007).

Irrespective of the status of the ultimately schematic level of linguistic knowledge, the usage-based view of L2 development as sketched above is essentially different from prevalent views in traditional SLA where development is seen as the gradual approximation of ever-systematic Interlanguage morpho-syntactic rules inclined towards those of a static end-point target language competence. This usage- and item-based approach differs from Interlanguage studies by not positing an end-point of learning, but also, and more importantly, by investigating the dynamic ontogenesis of interrelated usage patterns, rather than abstract rules thought to be broadly applicable across linguistic instantiations. In December 2006, in a special issue of *Applied Linguistics* focusing on emergentism and SLA, Larsen-Freeman espoused a view of language development as fundamentally different from that inherent in the traditional Interlanguage vein according to which target languages are viewed as stable and homogenous, Interlanguages are seen as purely cognitive and generally rule-governed, systematic, and predictable. Instead, Larsen-Freeman, advocating a view of language as both social and cognitive, sees language learning as more complex than being analyzable on the basis of predictable rule-learning; it is, she argues, to be understood as fundamentally dynamic, constant, and potentially unending (Larsen-freeman 1997, 2002, 2006; N. Ellis & Larsen-Freeman 2006). These general messages, fundamentally in congruence with the item-based emergentism suggested in UBL, are well taken and investigated empirically here.

#### Literature review: UBL in SLA.

To date, only a few studies have undertaken the task of testing the validity of applying the UBL path of learning to SLA which was originally suggested by N. Ellis (2002) as a default guide to investigating L2 development. Examining expression of future tense by 16 learners of English, Bardovi-Harlig (2002), in a response to N. Ellis, found two phenomena indicating that the path is only partially valid for SLA: 1) the role of formulas in initial development was found to be limited; and 2) the use of formulas was found not to diminish ontogenetically. The first phenomenon, Bardovi-Harlig concedes, may be due to the fact that initial formulaic use is too brief to be detected in her corpus. In terms of the second one, Bardovi-Harlig fails to acknowledge that the existence of the formulas in advanced stages of learning in fact supports the validity of the suggested path of acquisition insofar as it is connected with the cohabitation in the grammar hypothesis (Author in

press). Bardovi-Harlig does conclude, however, that N. Ellis' proposed path of acquisition presents a richer view of SLA development than a starting point that excludes formulas.

Bardovi-Harlig's final verdict which thus grants a place for UBL in SLA was corroborated in two studies so far, as Author and Colleague (2007) and Author (in press) successfully heeded N. Ellis' call for the application of the UBL path of learning to SLA. In the former study of the development of *do-negation* by a Mexican-Spanish speaking learner of English, Author and Colleague (2007) found positive evidence for the UBL path of acquisition, with the system emerging in acquisition characterized as the gradual abstraction of regularities that link expressions as constructions. *Do-negation* learning was found to be initially heavily reliant on one specific instantiation of the pattern, *I don't know*, with productivity gradually increasing as the underlying knowledge seemed to become increasingly abstract, as reflected in type and token frequencies. *I don't know* was also found to be stable throughout development suggesting its entrenchment as a MWE. These findings thus suggest that L2 learning is indeed item-based, that expression entrenchment is dependent on token frequency (as in the case of *I don't know*) and that more abstract pattern and construction learning is dependent on type frequency, as expected.

In a current longitudinal study on the same student, this time focusing on *can*-patterns, Author (in press) found no conclusive evidence that in L2 development, language knowledge becomes increasingly abstract; patterns could not be shown to be linked as abstract constructions in ontogenesis. Rather, this study found that certain MWEs were inextricably linked with certain interactional contexts, suggesting a locally contextualized view of L2 learning, and showing interaction to be a constant source of renewal for the individual linguistic repertoire. In other words, the study, proposing an emergentist outlook on development, suggested that language learning is an infinite process indistinguishable from language use. Productivity in language learning, the study suggested, seems to be guaranteed by a rich inventory of linked utterance schemas rather than the learning of increasingly complex combinatorial rules or increasingly schematized linguistic knowledge. In other words, it might not be the case for all kinds of linguistic patterns that their learning is a matter of abstracting regularities. Initially in development, however, the *can*-utterances were found to be dominated by a few concrete interrelated patterns, such as *I can write* and *can you write*, rendering language learning very much item-based.

Taking such concrete pattern instantiations as its starting point, the present study investigates creativity and routinisation in terms of how these patterns emerged into what they are at the time of utterance, and how they are interrelated at the time of the utterance. What this investigation will then present, hopefully, is a window onto how a linguistic inventory might look at a given point in time and how it came to be structured the way it is, taking to heart MacWhinney's (2006) encouragement to look for the roots of the emergent properties of language.

#### Data:

The data source for the present study is the Multimedia Adult English Learner Corpus (MAELC),<sup>23</sup> which consists of audio-visual recordings of classroom interaction in an English as a Second Language (ESL) context. The classrooms, in which the recordings were made, were equipped with video cameras and students were given wireless microphones on a rotational basis; the teacher also wore a microphone (Reder et al. 2003; Reder 2005). Consisting of recordings from July 2003 through July 2005 and September 2001 through February 2005, respectively, this is a longitudinal study of two students, Valerio and Carlos (pseudonyms), adult Mexican-Spanish speaking male learners of English, who were both judged to be successful learners (by standardized assessments and progress through the language school program). The final database of the inquiry consists of transcripts from approx. 70 sessions (each consisting of three hours of recordings, not all of which has been transcribed) in which Valerio and Carlos are either wearing a microphone or sitting next to someone wearing a microphone.

#### Data analysis and discussion:

Methodologically, the modus operandi is inspired by Lieven et al. (2003), who recorded a child, interacting with her mother, one hour daily over a six week period. The authors targeted the utterances made by the child in the final hour of recording and searched backwards in the database for "similar" utterances to trace the relationship with previously used linguistic material. Their basic finding was that 63% of their focal child's utterances were full verbatim repetitions of linguistic items used previously. Their focus, however, is on the 37% of her utterances that were novel and on the kinds of syntactic creativity needed by the child to produce the novel utterances. To investigate

---

<sup>23</sup> MAELC was compiled and is maintained at The National Labsite for Adult ESOL (known locally as the Lab School). The Lab School was supported, in part, by grant R309B6002 from the Institute for Education Science, U.S. Dept. of Education, to the National Center for the Study of Adult Learning and Literacy (NCSALL) and was a partnership between Portland State University and Portland Community College. My special thanks go out to Steve Reder and all the staff at The Lab School without whom my recent research would not have been possible.

this, they compared the target utterances with closest previous matches in the child's biography as a language user as it had been caught on tape. I will return to the content of this study as it becomes relevant during the course of the paper; for now please note that the procedure has been modified to fit my data, which cover a much longer period of time for both of my focal students. Fundamentally, however, I analyze spontaneous speech and trace the utterances in the same manner as described above. In more concrete terms, for Valerio and Carlos respectively I pinpointed 6 and 4 points in time distributed equally over their periods in the ESL class room. The reason for the difference between the students in this respect is related to data density; there is simply more data available for Valerio, and introducing this difference ensures a higher degree of comparability between the two students. For each student, the first recording point is approx. two months into their ESL curricula, making for a fair amount of searchable initial linguistic history, as it were. I then ended up with a number of target multi-word utterances (MWUs) for each focal point (roughly ranging from 70 – 130) which I then traced backwards in development in order to discern the interplay of routines and creativity in L2 development. Obviously both students employed single word utterances (yes, no, exactly, alright etc.) and tokens of reciprocity (uhuh, mhm etc.) which have been left out here because the focus of the investigation is on MWUs. Note at this stage that MWU is meant as a neutral term implying only the presence of several lexical items in an utterance as opposed to single-word utterances, whereas a multi-word expression (MWE; cf. definition of formula above) implies a possibly significantly recurring string of words.

The first step was to investigate the degree to which the individual linguistic inventories are characterized by the process of recycling. I therefore counted all audible, intelligible MWUs produced by the focal students at the designated focal points in the recording periods in order to investigate *the traceability* criterion; i.e., whether or not I could trace, in a straightforward manner, matches in previous use along the lines described above. The term "traceability" is chosen to denote the fact that even if I have been unable to trace given linguistic material in the data, it does not necessarily mean that it has not been produced or encountered before; the only certain thing is that it is not present in the data. Adhering strictly to the usage- and item-based principles, the tracing is a strenuous process of looking through the transcribed database for exact verbatim matches and recurring use of identical linguistic patterns over time. While the majority of the tracing was possible to on the basis of recurring items at the level of the "unit", i.e., word or construction level,

sometimes it was necessary to switch linguistic levels in the searches, as it were, and do queries based on morphological rather than lexical and/or constructional relations (e.g., in the case of *-ing*).

It transpired, however, that I could not easily quantify the MWUs while maintaining strict empirical authenticity. Due to the interactional setting of turn-taking (Sacks et al. 1974) in which the vast majority of the MWUs occurred, it seemed that the most optimal way of doing the counts was by attempting to operationalize the MWUs as multi-word turns-at-talk; i.e., to view interactional boundaries of turn-taking procedures as the most natural borders of the utterances. This was also feasible without loosening the stringency of the linguistic apparatus, because interactional turns may consist of several linguistic "chunks", turn-constructional units (TCUs; Sacks et al. 1974; Schegloff 1996), which are essentially compatible with the UBL view of linguistic units as form-meaning pairings (Fried and Östman 2005). Thus, in Tables 1 and 2 below, which present the results of the basic counts, the traceability criterion has been applied to all TCUs in the MWUs, whereas the quantification itself is based on MWUs as full turns-at-talk<sup>24</sup>.

Valerio	Summer 03	Fall 03	Winter 04	Spring 04	Summer 04	Summer 05	Total
Target MWUs	76	151	112	63	70	102	574
Untraceable MWUs	5 - 6,6 %	6 4 %	4 - 3,6 %	1 1,6 %	1 1,2 %	4 - 3,9 %	21 - 3,7 %
MWUs based on recycled TCUs	71 93,4 %	145 96 %	108 96,4 %	62 98,4 %	69 98,8 %	98 96,1 %	553 96,3 %

*Table 1: Emergent inventory, Valerio*

Carlos	Fall 01	Spring 02	Fall 03 – winter 04	Fall 04 – winter 05	Total MWUs
Target MWUs	43	80	121	120	364

<sup>24</sup> The compatibility between the basic unit in UBL and the TCU may not be as straightforward as implied here, but will be maintained nonetheless. Space does not permit me to enter the debate; see references cited as well as e.g Ford et al. (1996); Selting (2000, 2005); Ford (2004); Schegloff et al. (2002). Also, I have aspired to keep the turn-at-talk : MWU correlation as close as possible to a 1 : 1 correlation; however, maintaining this as a watertight taxonomy was impossible due to overlaps, long intra-turn pauses etc. in the classroom interactions.

Untraceable MWUs	2 4,7 %	2 2,5 %	5 4,1 %	3 2,5 %	12 – 3,3 %
MWUs based on recycled TCUs	41 95,3 %	78 97,5 %	116 95,9 %	117 97,5 %	352 – 96,7 %

Table 2: Emergent inventory, Carlos

Tables 1 and 2 give an overview of the linguistic inventories of the two focal students in terms of the degree to which their utterances employed are traceable, i.e., based on recycled linguistic material, such as MWEs, utterance schemas or verbatim repetitions of previous usage or material afforded (van Lier 2002) by the classroom interaction. In terms of overall relative numbers, the result of the counts displays remarkable similarity between the two students. Basically, and most importantly at this stage, we see that the vast majority of their utterances overall are based on recycled linguistic matter with only a tiny part of their utterances untraceable. It may also be observed for both Carlos and Valerio that there is a slight tendency for untraceable language patterns to decrease over time, albeit in a non-linear fashion. This is in part an artifact of the nature of the database, owing to the fact that the biographical linguistic inventories evolve, thus making for increasingly richer datasets to search through over time. However, the relatively small portion of untraceable utterances confirms the picture of their linguistic inventories as consisting largely of different kinds of recycled linguistic material throughout development. This rudimentary count, suggesting that the majority of the target utterances are derivable from previous usage, is in alignment with Author (in press), in which recurring utterance schemas were found to constitute the bulk of portable linguistic experience.

#### Categorizing the target utterances and the nature of recycled linguistic material<sup>25</sup>.

The next step in the investigation, then, is to create a taxonomy for categorizing the various utterances in terms of the extent to which they can be thought of as recycled. This is done on the basis of the nature of the items that are combined in use; i.e., no a priori categories are formed. The categories thus identified on the basis of concrete utterances in development are, respectively, untraceables, MWEs, afforded repetitions<sup>26</sup>, recycled utterance schemas, and various combinations

<sup>25</sup> *Anthology note: please see appendices 1-4 for detailed counts and utterance traces. Not included in original paper.*

<sup>26</sup> 'Afforded repetition' covers linguistic material afforded by either teaching material ('task-induced repetitions') or general classroom interaction ('interactional repetitions'). Please note that afforded material was only operationalized as such if it was found in the on-going interaction; i.e., if something was afforded interactionally, e.g., one week prior to

of previously used linguistic material. Due to the large numbers of traced utterances recounted in the tables above, however, it is outside the scope of this paper to present a total overview of all target utterances, almost 1,000 MWUs, and their respective derivations. Apart from the more detailed count displayed in the Appendix which displays the results of the counts of these categories for recording periods 1 and 4 for Carlos, and 1 and 6 for Valerio, representative exemplars will have to do. The point of this section is, therefore, to briefly explain and exemplify the categories, which in the interest of readability is underlined in the following.

Language patterns which seemed to have no concrete predecessor in terms of the concrete items and / or patterns involved were defined as untraceables. As can be inferred from the tables above as well as the Appendix this did not apply to many utterances and this category will not be dealt with further here.

The rest of the inventories consists of full or partial repetitions. The category of recycled linguistic matter of least importance for the present purposes, and also the most complex and hard to define category, were those utterances which consisted of a combination of previously used material, but were not empirically substantiated to be sanctioned by utterance schema-like patterns. This is a very small group of utterances, as may be inferred by the Appendix, and apart from the following example, Valerio's use of the expression *just working*, this category will not be dealt with further here:

Extract 1<sup>27</sup>:

(August 12 2003)

1 Teacher so how're you doing mister?

2 Valerio → just working ((waves hand))

3 ()

4 Teacher yeah [ok

5 Valerio [xxx

---

deployment of a given target utterance, it was not counted as an affordance. Examples will be given below of material afforded interactionally. The notion of 'affordance' owes to van Lier (e.g. 1996, 2002)

<sup>27</sup> Transcript conventions: xxx = inaudible; [] = overlap; () = pause; : = prolonged sound; (...) = transcriber's comments. ?: rising intonation; . = falling intonation; U = unidentified student. Arrow = the target expressions. -- = break.

*Just* has been used before in the expression *just for the years*; *work* among other places in the expression *quit my work*; and the progressive has been used in different contexts, among them being *my wife she shopping for me now*. So all components of the expression, as it were, have at this point been in use before. It could be argued that there is a *just x* schema; however, *just for the years* is the only other instantiation of *just* in the data, and that particular expression is more likely to be an instantiation of the frequent *for NP* pattern with *just* added on to it.

In terms of the next category in the taxonomy, full repetitions, which is a bit more consequential one for the present purposes, we note that these are found in two varieties; 1) MWEs, which are operationalized as recurring strings of words with a relatively coherent and constant communicative purpose (Author and Colleague 2007; Author in press); and 2) afforded repetitions, which were found to be verbatim repetitions of something that had somehow been afforded (van Lier 2002) by the immediate classroom interaction<sup>28</sup>. The extracts below display examples of MWE (extracts 2a+b) and afforded repetitions (extract 3).

Extract 2a:

(Oct 04 2001)

- 1 Teacher    okay the date () what's today
- 2 Multiple    today Tuesday
- 3 Carlos → ((raises hand)) I can write
- 4 Teacher    ((walks toward the board, holds pen high, waves pen))
- 5 Carlos → I can write ((reaches for pen))
- 6 Teacher    you always write Carlos
- 7 Carlos     heh
- 8 Teacher    okay ((hands another pen to another student))

Extract 2b

---

<sup>28</sup> A third category might be identified as a verbatim repetition that is neither 1) nor 2), but merely something that is repeated. This question of defining MWEs as opposed to other less crucial repetitions is a tricky one having to do with both aspects of frequency and interactional anchoring and will not be solved here, where the focus is on utterance schemas.



(May 17 2002)

- 1 Teacher okay () now I need two students () here's one  
2 ()  
3 Teacher where's the other one  
4 Carlos → ((gets up)) I don't know () but I can write () write ((reaches for Teacher's  
5 pen))  
6 Teacher ((gives pen to another student))  
7 Carlos oh he? ((points to student, goes back to desk))

Extracts 2a and 2b, transcripts from two similar classroom situations, occurring 8 months apart, clearly show the stable meaning and interactional anchoring of the MWE *I can write*, as Carlos' deployment of this MWE is strictly tied to situations where the teacher wants something written on the public board in the classroom (see Author in press for a full account of the locally contextualized nature of this MWE).

An example of afforded repetition, underlining the situated nature of the affordance, comes from Valerio:

Extract 3

(August 12 2003)

- 1 Valerio it's mm the () no is different ((shaking head)) for [red  
2 Teacher [it's not different?  
3 Valerio → it's no differ[ent ((shakes head))  
4 Teacher [it's not different.  
5 Valerio no ((shakes head))  
6 Teacher it's just the size is different?  
7 Valerio mhm: ((nods))  
8 Teacher okay

In this interaction, Valerio and the teacher are talking about differences in fruits between Mexico and the US. Valerio expresses, in line 1, the similarity between different kinds of melon in the two countries. The teacher seems to want to direct Valerio's attention to the non-native-likeness of his negated pattern and in line 3, the instantiated afforded repetition, Valerio appropriates the teacher's utterance from line 2.

The two kinds of verbatim repetitions thus identified seem to be important ingredients in the L2 linguistic inventory in different ways; the afforded repetitions play a role in terms of the interactional activities in this classroom, but less so, it would seem, in terms of long-term retention of linguistic units. The MWEs, on the other hand, constitute a qualitatively important ingredient both in terms of classroom interaction and long-term issues of linguistic sedimentation, as show in extracts 2a-b above. However, unlike the focal child in the study by Lieven et al. (2003), my focal students do not seem to operate primarily on such verbatim MWEs. Whereas Lieven et al. found that 63% of their focal child's utterances were recycled in their entirety, no such explicit count was undertaken for these data. However, the fact that MWEs are an important yet numerically inferior kind of linguistic item in the emergent linguistic inventories of my focal students is evident in the Appendix which displays a more elaborate count of the various expressions deployed by the two students at the end of recording periods 1 and 4 for Carlos and 1 and 6 for Valerio. Afforded repetitions decrease dramatically for both students, and MWEs, while constituting a considerable part of the recycled linguistic matter, does not seem to represent the bulk of linguistic experience.

Rather, the most important category, constituting the bulk of utterances produced, are recycled utterance schemas, which are partial repetitions, brought in by extra material, e.g., something that was afforded by the immediately accessible interactional environment, a lexical substitution, or an add-on as in the example with *just for the years* above. The stuff of traceability, in other words, is not necessarily based on actual verbatim repetitions but on recurring and seemingly broadly applicable patterns, utterance schemas, consisting of a stable part (i.e., that which is reused) and a flexible part (an open slot in the pattern). The question to be answered is therefore how this bulk of linguistic experience, the portable sediments of linguistic interaction in the form of recycled utterance schemas, is organized in Valerio's and Carlos' respective linguistic inventories. Illustrative examples from the data will be given below in section 4.2. Section 4.3 will then investigate the emergence of such utterance schemas.

### Tracing the target utterances and identifying the utterance schemas:

The point of departure for this next undertaking is based on Lieven et al. (2003) whose study focus was to define the syntactic operations required by the child to produce the 37% of her utterances that are novel. They identified five syntactic operations: "substitution", "add-on", "drop", "insert", and "rearrange". If for example a target utterance was *I got the butter* and the closest match was recognized as *I got the door*, the syntactic operation identified to be required for producing the novel utterance was "substitute" (the object). Add-on, drop, and insert refer to items that in the target utterances have been added, dropped, or inserted as compared with previous closest matches (e.g. *let's move it around* from *let's move it*; *and horse* from *and a horse*; and *have you finished with your book* from *have you finished your book*, respectively). The final operation, rearrange, is thought to be applied in utterances that consist of the same items as a previous utterance but employed in a different order, as exemplified by the target utterance *away it goes* from the traced match *it goes away*.

As will become clear in the following, the five syntactic operations identified by Lieven et al., while valid to some extent also in terms of the present data, quickly seemed insufficient to adequately, exhaustively, and systematically account for the ways new utterances emerged from previous ones in terms of the L2 classroom. In Lieven et al., 74% of all novel utterances required only one operation (and most of them "substitution"). The target utterances in my data seem to require much more than one operation to be arrived at when compared to their previous closest matches. Data examples from Valerio's first 10 weeks of class attendance, corresponding to recording period 1, have been chosen to show how even in initial development, the number and kinds of syntactic operations required are too complex to be captured in any straightforward manner. The first two data examples are backed by extracts from the database because aspects of interaction seem to have a direct impact on the utterances used.

Target utterance: *I like the movies the dancing*

→ closest match: **I like** gold (schema **I like x**; operation: substitution)

In terms of the specific lexical substitution, *dance* has been used before, whereas *movie* has not. It is argued, however, that this specific lexical item is nonetheless recycled because it is afforded by the immediate interactional context, as shown in extract 4.



Target utterance: *here is big*,

→ closest utterance match *is good no good* (schema: **It's adj.**<sup>29</sup> Operation: substitution + add-on)

To arrive at the target utterance from the closest match, Valerio not only substitutes the adjective, he also adds an item to the utterance schema. The lexical item added to the schema is *here* which underlines the locative nature of the intended meaning; "Vegetables taste the same in Mexico and the US but in the US they are big". The item itself, *here*, is not only traceable to previous uses, but may also be argued to be emergent, adopted from the environment as it is afforded by the ongoing interaction; lines 2 and 5 in the extract below:

Extract 5

(12 Aug 2003)

- 1 Teacher ((instructing)) okay it says completing sentences with but () add a logical ending  
2 to each of these sentences () number one () in my country we eat a lot of fresh fruit  
3 but here we eat  
4 U xxx  
5 Teacher you eat fresh fruit here  
6 U xx[x  
7 Teacher [we eat  
8 Multiple xx[x  
9 Teacher [eat  
10 Valerio n[o:  
11 Teacher [fresh fruit? () banana () apple oranges when you buy them in outside in the  
12 mark at the market () what is the different what is a different way of eating fruit ()  
13 when they are in  
14 Valer → big here is big () heh  
15 Teacher here they are bi:g

---

<sup>29</sup>Note the conflation of *is x* and *it's x* schemas due to audibility factors; I have relied on the context to identify the schema as 'dummy subject' rather than 'regular copula'.

The following examples from the data are not backed by extracts. What unites them is that they are very difficult to handle in terms of syntactic operations. To exemplify, two different utterances which are both derivable, as it were, from the same closest match, will be discussed.

Target utterances: 1) *No is melon*; 2) *No is different for red*

→ no is for profession

In terms of 1), at least two operations have been required, namely substitution and drop, for *for profession* to become *melon*. The problematic thing here, however, is that the meanings of the two are different. *No is melon* means something along the lines of *it's not melon*, it referring to "the English word", and the other one means *it doesn't apply to her profession* or something to that effect. While the nature of syntactic operations says nothing about this issue, it is not a problem as regarded from a construction grammar perspective; it would go under the heading of constructional polysemy (Taylor 2002), and we would simply posit multiple meanings for the *no is*-pattern. To complicate matters further, the traced match is likely to be an utterance schema combination, consisting of two patterns, *no is x + for NP*. The *no is x*-pattern, in any case, seems to be a pattern in its own right. One instantiation, *no always is* (with the subject predicate presupposed), seems to have been arrived at by inserting an adverb into the schema. Insertion might also be in operation in 2) above for which multiple operations are needed. If we can substitute *for red* for *for profession* we also need to insert *different*. However, the negated parts are *different* and *for profession*, respectively, so those are the two items that are substitutable. In order for this to be valid, a further operation is required to arrive at the target utterance, namely in the form of *for red*. To make matters more intricate, this is an instantiation of *for NP* consisting of a combination of items which have only been used separately in other previous contexts, which means that there is also a substitution to deal with in terms of the *for NP* pattern.

In other words, to the extent that the syntactic operations are easily delineated for each target utterance, the question of number and kinds of operations required presents a nagging problem inasmuch as already at this early point in development, the processes involved in producing the utterances seem very complex. The final example in this section underlines this high level of complexity:

Target utterance: *I am forget*

→ I am no forget

→ I am go, I am visit, I no understand.

This chain shows the tracing of another pattern, the non-native like copula-verb pattern. *I am forget* has its closest match in its own negated counterpart, which again is traceable to instantiations of the pattern per se and an instantiation of another negated utterance, supporting the assumption that *no-negation* is also a pattern (Author & Colleague 2007). However, if the *no*-negation is an utterance schema in its own right, which the data seem to support, it becomes more likely that *I am forget* is the result of a substitution from *I am Verb* rather than the result of a syntactic drop of the negation particle from its own negated counterpart. To complicate matters further, the target utterance above is taken from a larger utterance: *I am forget for the last name and first name you*. What we have, then, is the combination of two utterance schemas, *copula + verb* and *for NP*, with an additive *and* inserted into and with *you* added on to the *for NP* schema. The actual number of required syntactic operations seems to be quite uncertain; however, the utterance itself, in terms of previously employed utterance schemas, remains straightforwardly traceable.

With the evidenced plurality of syntactic operations required to get to target utterances from previous matches already at this early stage in ontogenesis (remember, the utterances are all from Valerio's first ten weeks in the classroom), it seems impossible to describe development of the interplay among utterance schemas, lexical items and MWEs in terms of number and kinds of syntactic operations along the lines defined by Lieven et al. (2003). Consequently, such syntactic operations seem to be invalid as tokens of L2 development as conceptualized here. This may be suggestive of the incommensurability of L1 and L2 learning research; Lieven et al. are specifically searching for evidence of usage-based syntactic operations in a child, whereas I am looking for more long-term development, evidence of L2 learning, in two adults. In any case, development evidenced in the data seems too complex and non-linear, the utterances perhaps not learnt by way of formal causality, to conform to strict terms of syntactic operations, thus confirming the views espoused by Ellis and Larsen-Freeman (2006) as recounted above. The kinds of syntactic operations, to the extent that they are at all psycholinguistically valid, are too complex and too varied already in this preliminary developmental phase, for L2 development to be framed upon issues of syntactic processing; rather, it is the tracing of the emergence of utterance schemas similar

to the ones presented in the examples above – the investigation of how they develop – that will shed light on how L2 learning takes place along a usage-based path of routinisation and productivity.

As reflected in the tables above in the numbers conveying the relationship between total utterances and recycled utterance material, most notably utterance schemas, the linguistic inventories of my focal students are at all times dominantly characterized by a set of recycled patterns. So far, however, we have said nothing about how these patterns are interrelated or how they develop, or if the same patterns are dominant at all times. Below, we shall have a look at how my focal students' inventories have been shaped by investigating in closer detail the development of certain focal utterance schemas and MWEs. By doing so, we investigate “the ‘natural history’ of particular utterances” (Lieven et al. 2003: 335) in order to get at how the interrelationship is among recycled material and new (mostly lexical) additions, both cross-sectionally and longitudinally.

#### Emergence of utterance schemas.

Having validated the view of the inventory as consisting primarily of utterance schemas and MWEs through the empirical demonstration that an overwhelming majority of the focal students' utterances are in fact "built" from such items at all times in ontogenesis, we now turn to investigate, not so much syntactic creativity and psycholinguistic operations, but the interrelationship among utterance schemas, MWEs, and creativity. The Appendix presents an overview of the counts at the end of recording periods 1 and 6, and 1 and 4, respectively for Valerio and Carlos. While both inventories have grown in absolute numbers over time in terms of both MWEs and utterance schemas, the interesting observation lies, yet again, behind the numbers, in the possible interrelationships among the utterance schemas in development. An example is the *it's x*-schema which seems to have evolved into a wider variety of interrelated schemas, such as *it's more x*, *because it's x*, *this is x* etc.. Presently, then, we seek to investigate at what point in Valerio's development the schema *is/it's more x* emerged, perhaps from the preceding schema *is/it's x*. Towards the end of recording period 6, the following interaction takes place. The students are doing a "free movement" task, which is fairly common praxis in this classroom, the students moving around freely engaging relatively briefly in cued, serial dyadic interactions (Hellermann and Cole forthc.). The cue in this particular case consists of the students having been equipped with little cards stating a "comparison". The female peer has a card saying "living in a house or in an apartment":



Extract 6

(July 28 2005)

- 1 U female    you  
2 Valerio → eih ((reads her card)) living in a house is: () is more rel: no not relax uh it's more  
3            u:hm pf () hi:h () I don't mean relax what is ((moves from side to side))  
4 U male      comfor[table  
5 Juan                [comfor[teh  
6 Valerio                [comforte:hble  
7 U female    comfortable?  
8 Valerio      comfortable ((nodding)) to to the apartment to live in apartment

While Valerio's first turn here seems to consist solely of recycled linguistic material, hence presenting an interesting example of how combinability is about putting together chunks rather than lexical items as separate islands, this is not my primary interest at this stage. Rather, what I am after is the "natural history" of the utterance schema. Because the schema itself is not among Valerio's initial resources, it is necessary to go by the concrete item *more* in order to trace the development of the schema. This reveals, for the first 7 months of his time in ESL class, 17 instantiations and a scattered use of the item *more* in various contexts. There are, however, two dominant uses of the item, namely a *no more x*-pattern, and a *one more x*-pattern, emerging in recording periods 2 and 3 respectively, accounting for 10 of the 17 instantiations. At the same time, we have the development of the highly useful and frequent *it's x*-pattern. Without going into detail on Valerio's varied uses of the pattern, suffice it to say that most often he uses it straightforwardly in *it's + adjective*, e.g. *it's good, it's okay, it's easy, it's difficult*. Looking at *more* and *it's adj* in parallel emergence from the broad etic perspective it seems to be a matter of experience before the two are combined. The turning point in this respect seems to be June 25<sup>th</sup> 2004, approx. one year into Valerio's biography as language learner in this classroom:

Extract 7

(June 25 2004)

- 1 Eric            ((laughs)) () I came you\_ you\_ you came to America y\_ two years ago

2 because you think what different is it in your country and and here  
3 Valeri → here? () the culture and um () and the: () people and the city () all the () is  
4 more uh in Portland is more um () easy for you walking in the night () in  
5 Mexico it's it's very very difficult  
6 Eric huh?  
7 Valerio it's ver\_ it's big Mexico City ((nodding)) but for Mexico all Mexico is  
8 different () is different than () you check the other map ((pointing))  
9 --  
10 Valerio here you pay seven dollars per hour bu:t in Mexico you pay seven dollars  
11 for () two days ((motions two))  
12 Eric two days  
13 Valerio or  
14 Eric one day three hours work  
15 Valerio you eight hours or more  
16 Eric they pay for more or pay in two [days  
17 Valerio [maybe eight  
18 Eric twenty hours\_ twenty hours seven dollars ten dollars you mean  
19 Valerio maybe one dollar per hour in Mexico maybe.  
20 Eric maybe heh heh h[eh  
21 Valeri → [but no is\_ is good is (+) little it's\_ it's all the people  
22 Mexico is here because it's eh the work is easy and the money is more for  
23 Mexico  
24 Eric but if you put uh to buy something or () I mean there () is is e\_  
25 expensive or or or the same  
26 Valerio if immigrant () normally () the \*mayori the \*mayori people live here is the  
27 () immigrant no is legal ((\*majority))  
28 Eric mhm heh heh heh heh  
29 Valerio but it's little bit people live here and it's legal ((nods))  
30 Eric ((nods))  
31 Valeri → and so the majority men (+) men is live here is coming here because it's  
32 more uh () it's more easy for pay the money in Mexico for the family for  
33 the children () it's\_ it's more easy here

- 34 Eric hm
- 35 Valerio you live here and you: you put the money for your wife [and xxx]
- 36 Eric [mhm ]
- 37 Valerio → it's more money (+) it's\_ it's more easy

This interaction shows the coming together of the *it's x* and *it's more x*-patterns. They have not emerged from the same source with the latter being a much later addition to the inventory than the former. Creativity, then, is a matter of "discovering" new schemas, building new schematic routines on the basis of previous ones. Moreover, if this interaction is in fact the ontogenetical source of the *it's more x*-schema, it shows emergence to be locally contextualized and item-based; *it's more easy* (even in close collaboration, as it were, with related schemas) recurs here and forms the backbone of schematic development (Author and Colleague 2007). This, then, is where the new schema seems to emerge. Such evidence lends empirical credence to the idea that use and learning are inextricably intertwined (Firth and Wagner 2007); they cannot be divided into "interaction" and "acquisition" with the intervening notions of "input", "intake", and "noticing" as usually suggested in the input-interaction framework (e.g., Gass 2003). The fact that the present approach distinguishes itself from the input-interaction framework is also evident in the case of Extract 3 above where the ideas of interaction, input, and acquisition cannot be validated in the data; Valerio did not "notice" or "take in" the native-like variant of the negated utterance at that particular point in time but kept using the *no is*-pattern in a variety of contexts after this interaction.

A similar, yet different, form of development is seen in Carlos' inventory. It is similar to the extent that we seemingly have a schema emerging from previous usage; it is different to the extent that this time, rather than emerging into a "bigger" schema in terms of volume, we see development as going from the *how many x*-schema, to the smaller, highly general *many x*-schema. Carlos' first recorded use of *how many x* is *how many you have* ( ) *how many how many* in a context in which Carlos and his partner are discussing pair work results; a context which recurs with the same MWE *how many you have* 6 months later. Incidentally, *how many you have*, not only instrumental in the emergence of the *how many x*-pattern, also represents the first recorded spontaneous use of *you have* by Carlos<sup>30</sup>. While this is not crucial to the pattern under investigation here, it is an important discovery insofar as it shows the interdependence of the patterns in development. Meanwhile,

---

<sup>30</sup> He does use it once earlier in a session in which *do you have NP* is the targeted form of the practicing.

approx 4½ months after the first recorded use of *how many*, this interaction takes place in which *how many* is used 7 times in conjunction with *years* and *have*<sup>31</sup> (extracts 8a-8c):

Extract 8a

(March 05 2002)

- 1 Carlos → how many how many grandchildren do you have?
- 2 Alexander I fourteen grandchildren
- 3 Carlos eh?
- 4 Alexander fourteen
- 5 Carlos fourteen?!
- 6 Alexander heh [heh heh
- 7 Carlos [heh heh heh

Extract 8b

(March 05 2002)

- 1 Carlos → heh hah how many how many: () how many years old have ehm e:h your
- 2 las:t children
- 3 ()
- 4 Alexander how many children?
- 5 Carlos → no how many years old () years
- 6 Alexander how many years:
- 7 Carlos years old
- 8 Alexander oh

Extract 8c

(March 05 2002)

- 1 Alexander grandmother () you have grandmother?
- 2 ()

---

<sup>31</sup> The *how many years do you have* use is probably L1 transfer.

- 3 Carlos     yeah the: mother () of my: of my mother  
 4 Alexander grandmother  
 5 Carlos     yeah [my grandmother  
 6 Alexander     [your grandmother huh?  
 7 Carlos     ((nods)) my grandmother  
 8 Alexander e[:h  
 9 Carlos     [she's uhm  
 10 Alexander how many years old?  
 11 Carlos → uh uh I don't know how many years she has maybe she has eighty:: eight  
 12             eighty something like that  
 13 Alexander oh oh oh  
 14 Carlos     yeah

For this interaction, the students had been put together in pairs and given the task of asking each other questions. The extracts, transcripts from the same on-going interaction in the classroom, show a dense frequency of use of the targeted item, supporting the idea that the learning of the pattern, the utterance schema, is item-based. After this session, use of the pattern becomes more scattered across interactions in the classroom with a more varied use in the open slot position. This is exemplified by extracts 8d-e, recorded almost two years after 8a-c:

Extract 8d

(January 20<sup>th</sup> 2004)

- 1 Nadja             and then we'll have a: Chinese folk dance dragon dran dragon dance () do you  
 2             know that? () ((points to newspaper)) dragon dance  
 3 Carlos →     hh dragon dance yeah ((points to same spot in newspaper)) () it's () it's about ehm  
 4             how many people in e:h () in the the dragon ((gesturing)) because it's one person  
 5             in fro[:nt  
 6 Nadja             [uhuh uhuh [yeah  
 7 Carlos                     [uhm () three other persons in the middle  
 9 Nadja             w[e have xxx ((stretches arms backwards))  
 10 Carlos             [with the: ((stretches arm backwards))

- 11 Nadja uhuh ((nods))  
 12 Carlos → how many people in the the dragon () how many people  
 13 Nadja uh xxx  
 14 ((both orient to teacher))  
 15 Nadja can be: can be: I don't know () five maybe () five at least ((nods))  
 16 Carlos yeah maybe ((nodding))

Extract 8e:

(January 20<sup>th</sup> 2004)

- 1 Carlos you have a good job? or no? xxx  
 2 ()  
 3 Emilio I'm working  
 4 Carlos yeah you working [there?  
 5 Emilio [finally  
 6 Carlos yeah? heh heh heh heh heh  
 7 ()  
 8 Emilio it's [part part-time  
 9 Carlos [how m  
 10 Carlos → part-time yeah how many days a week  
 11 Emilio hm?  
 12 Carlos → how many days a week [do you work  
 13 Emilio [three days  
 14 Carlos three days?  
 15 Emilio yes five hours per day

The interactions in extract 8d-e are not similar. 8d, from a non-task interaction, is in the beginning of class; as a form of pre-activity, Carlos's partner is reading a newspaper and Carlos, interested in the article which is about Chinese new year celebrations, starts asking questions about what she is reading. 8e is more closely related to the task in which Emilio, Carlos and a third student are asking each other questions about what they did before coming to Portland, where they used to live in their home countries and other things pertaining to their respective backgrounds. The target utterance is

made relevant through Carlos' insistence to learn more about one of the issues he asks Emilio about, namely the job he used to have when he lived in LA.

Carlos' use of the *how many x*-schema seems to have emerged from contexts of comparing task results with alternate partners in the classroom. Biographically he then moved on to employ it in a more general sense to ask interlocutors about different things pertaining to their lives, past and present, an ability which seems to have come from a specific interaction, extracted in 8a-c above. These interactionally different situations in which he comes to deploy the pattern over time, underlines the portability of utterance schemas and the usefulness they represent. More fixed patterns may be fixed to certain interactional environments, as was shown in the case of Carlos' use of *I can write*, but the characteristic of being portable across contextual boundaries (Larsen-Freeman 2004) seems to be the prerogative of utterance schemas.

Concluding this exploration of Carlos' deployment of *how many x* and *many x*, Extract 8f below displays Carlos' first spontaneous use of *many x*<sup>32</sup>. The teacher directs her question to the whole classroom and after a brief silence Carlos self-selects as the next speaker to answer her question relevantly *many years*. (The teacher turns the situation into a focus on how to use a "full sentence", *I've been here for x years*, but that is beside the point here and hence excluded from the transcript).

Extract 8f

(January 20<sup>th</sup> 2004)

1 Teacher    how long have you been in Portland?

2            ()

3 Carlos →   many years heh heh heh

In other words, Carlos' first *many x* use includes (if it is not directly dependent upon) the use of a recycled noun in the utterance; a recycled noun that had even been recycled in the schema speculated here to be the predecessor of the current one. Thus, there is a very tangible link between the two schemas investigated here, namely the common use of the item *years*. What we have, then,

---

<sup>32</sup> *I would like to make many things* is an earlier pre-formed usage (approx. 3 months prior to this one) which Carlos seems to have written down before uttering.

is a MWE, *how many you have*, which is expanded to *sanction*, first, *how many x (do) you have*, secondly *how many x*, and ultimately *many x*, with a heavy co-dependence on shared lexical material between the usage patterns, as it were.

Another form of creativity development is found in terms of open slot variance (Author and Colleague 2007). An illustrative example of this process can be demonstrated in these data as well through a relatively simple type-token count and is found in Valerio's *in + location*-schema. Picking three intervals over two years we see the following development in creativity:

<i>In + location</i>	Pattern Usage			Afforded repetitions
	Token	Type	Ratio	
July - Nov 2003	26	12	0.46	7 – 26,9 %
Jan – Apr 2004	38	24	0.63	10 – 26,3 %
June – July 2005	31	25	0.81	0 – 0,0 %

Table 3: emergent creativity in Valerio's *in + location* schema.

Table 3 displays an increasing type-token ratio, suggesting increasing variance in the open slot in the schema. Furthermore, the drop in afforded repetition ratio supports an increased ability to creatively use the schema. The *in-location* schema does not only serve as illustrative of this dimension of creativity in language learning; it also supports the idea of the process of routinisation of the schema to be item-based. This is indirectly shown in the type-token ratios in the schema. Dwelling for a second on the 2003 ratio, we note a relatively low creativity. This would point in the direction of one or more frequent specific items, perhaps MWEs, in use for this schema. And that is exactly what we find. 20 of the 26 uses in 2003 involve a country or the word "country", and 6 of these may be identified as the MWE *in my country* (a MWE which is also in use throughout development, the final recorded use of the expression being June 2005). Especially one interaction seems to be pivotal in the emergence of the *in + location* pattern:

Extract 9

(August 12 2003)

1 Valerio → teacher in the[:h



2 Dom [yeah

3 Valerio → sentence for ((starts reading)) in my country most people go home for lunch at

4 ((stops reading)) () no is at noon is at afternoon

5 Dom at noon uhuh () is it at no[on or is it in: ] the afternoon

6 Valerio → [in my country is]

7 Valerio → in my country it's afehternoon

8 Dom afternoon [what time

9 Valerio [the united spate the united estate people he like the noun hn hmf

10 Dom they li[ke no]on ok at noon

11 Valerio [like ]

12 Dom but in your country [it's in

13 Valerio → [in my [country

14 Dom [i:n

15 Valerio [i:n the afternoon

16 Dom [xxx

17 Dom afternoon ((writing on board))

And approx. one month later, we may observe the pattern to be routinized, perhaps even beating in competition the much more frequent for NP-pattern, when used to talk about locations.

Extract 10

(September 23 2003)

1 Valerio in Cheena what is

2 Jing in [Ch\_

3 Valerio [de:h

4 Jing in China

5 Valerio ((nods)) in Chin[a

6 Jing [in China () uh summer is a uh is a raining raining xxx

7 Valerio a::h (nods)

8 Jing uh there are many rain many rain in summer

9 Valerio alright ((nods))

- 10 Jing but in the autumn and eh winter () eh the rain is xxx
- 11 Valerio →((nods)) alright () in Mexico for my country [it's
- 12 Jing [Mexico is very hot
- 13 Valerio →in Mexico no: for for my country in my country i:s this is similar
- 14 Jing similar
- 15 Valerio the raini:n[g and little summer]
- 16 Jing [oh raining ] o:h

Summing up at this point, language portability seems to be a matter of utterance schemas and, to a lesser extent, MWEs. These do play a prominent role interactionally and acquisitionally as the item-based process of learning gets underway. This process, however, is essentially dependent on locally contextualized expressions whose importance seems to be informed by interactional requirements to such an extent that if the interactional need to deploy them wanes so will their presence in the data. The linguistic matter which is most easily transported across the borders of these interactional requirements is a bit more generic in its deployment potential. However, as we have seen explicitly in the case of the coming together of *it's x* and *it's more x*, such schemas are not learnt in an interactional vacuum; rather, use, acquisition, and interaction is a trinity – perhaps the holy one of L2 learning – which is fruitless to disintegrate. We simply cannot tell them apart. Thus, creativity development can be captured in terms of utterance schema emergence; new schemas emerging from previous ones. This finding is in alignment with and further adds to the insights generated by Author and Colleague (2007) and Author (in press).

New schema emergence and variety in the filling of an open schematic slot thus seem to be the primary tokens of increased creativity in L2 development, at least as far as these data are concerned. However, another token of L2 learning, perhaps, is found in terms of combinability as an increase in number of intra-turn utterance schema operations, as it were. The combinability itself does not seem to require increasingly complex structural grasp of the target language. Two target utterances from each of the focal students' inventories serve as illustration:

➤ Valerio, Period 1:

Target utterance: *teacher in this sentence for* (starts reading) “in my country most people go home for lunch at” *no is at noon is at afternoon*

→ teacher (add-on / invocation) // in + location // afforded material (reading) // no is x // at + time // is x // at + time. *Teacher* is a frequently employed utterance-initial item used by Valerio to establish reciprocity, *in + location* is a proposed schema, initially almost exclusively coupled with the recurring item *in my country*, *no is x* and *is x* are both repeated utterance schemas early in ontogenesis, the latter already having been dealt with above. The two *at + time* instantiations are both afforded by the environment so the schematicity here is speculation. The important thing to note is that the utterance is a combination of previously used utterance schemas and interactional affordances.

➤ Valerio, Period 6:

Target utterance: *is it's perfect for because it's cheap and they have all things you need for the first time for one baby time it's perfect*

→ it's x // for // because it's x // and // x have // all things // you need // for NP // for NP // it's x. *All things* and *you need* are both verbatim repetitions, perhaps MWEs, as is *it's cheap*. *One baby time*, consisting of previously used material, seems to be creative (and does not seem to conform to a purist view of grammaticality). Both Valerio's utterances here are combinations of previously used elements, recycled utterance schemas, and verbatim repetitions. The one from period 6 seems to be characterized by more intra-turn schema usage.

➤ Carlos, Period 1:

target utterance: *Because my cousin wife they go in the California*

→ because x verb (uncertain number of substitutions) go + direction / in + location (substitution)

➤ Carlos, Period 4:

target utterance: *Yeah because I go to the\_ to dance at f's hideaway and many people from the many states they dance very well like Japanese people and Chinese people*

→ Yeah because // I go to the // to dance // at f's hideaway // and // many people // from many states // they dance very well // like Japanese and Chinese people //

→ yeah x // because I verb // go + infinitive // at F's Hideaway // and // many x // from + location // they dance // very adjective // like x

Carlos' utterances, like Valerio's, consist of complex assortments of recycled lexical items, patterns, and MWEs. As soon as we abstract away from the concrete utterance, however, things become speculative. It is, for example, difficult to determine if a repetition is a bona fide MWE or merely an arbitrarily repeated item. The kind of dynamic interplay among recycled lexical items, patterns (utterance schemas) and MWEs displayed by these four utterances, however, is found throughout development, but with an inclination towards an increase in intra-turn schema operations, as it were. The kinds of operations deployed do not seem to undergo fundamental changes over time, whereas the number of such operations seems to increase. This might prove a fruitful point of study for further UBL research in SLA.

These various considerations result in a picture of the emergent linguistic inventory as one consisting of moveable, manipulable, and flexible partially schematic, partially specific patterns and MWEs. Demonstrating an increasingly abstract underlying representation as a form of L2 learning seems to be rare. This does not mean that combinability is no longer worthwhile; it does mean, however, that we need to rethink it. Pawley and Syder (1983), in their seminal paper in which they addressed two puzzles for linguistic theory, introducing the lexicalized sentence stem (LSS) and inspiring much future research in formulaic language for SLA (e.g., Nattinger and DeCarrico 1992; Moon 1997; Nesselhauf 2005), tried to get at the fact that we as native speakers do not exploit the potential of grammatical combinability, "we do not standardly combine our utterances all the way from minimal items to complete utterances" (Harder 2001: 234). This is parallel to what we see enforced here in terms of both (1) processes and (2) goals of L2 learning. In terms of (1), learners do not pass through stages of syntactic development on their predestined way to approximating near-native mastery of a target language competence (whose stability is a fiction, anyway), and in terms of (2), utterance schemas (roughly parallel to Pawley and Syder's LSSs) are the linguistic stuff of L2 learning rather than rules of syntax. This is the step we are now able to take as we can empirically substantiate a justification of ruling out syntactic combinability as the stuff of learning and instead put in item-based utterance schemas as the linguistic material to learn. The next step is to investigate more closely the reflexive nature of participation in the classroom interactions and the

use and learning of utterance schemas and, consequently, L2 development (Author forthc. a., forthc. b).

### Conclusion

Understanding the object of research in longitudinal L2 studies as a hybrid between locally applied usage patterns and application of the same and related usage patterns over time, this paper has investigated ontogenetic development in terms of an empirically grounded, emergent grammar, consisting of units of spontaneously occurring language use. This has resulted in a conceptualisation of emergent creativity as building on recycled linguistic matter in the form of MWEs and utterance schemas.

Exploring in detail the relationship between routines and creativity in linguistic development in a second language, this paper has empirically substantiated L2 development to be too complex and non-linear, the various utterances perhaps not learnt by way of formal causality, to be exhaustively captured by strict terms of syntactic operations which in the study by Lieven et al. had proven to be a very fruitful approach to early L1 development. We have seen here that L2 development is not that easily captured, that the linguistic inventory (in a L2) seems to be a structured set of utterance schemas, patterns that are more or less lexically specific, but which overwhelmingly seem to stem from very specific patterns themselves, showing language development to be item-based. The language used by the two focal students here seems to be comprised of different kinds of lexically specific patterns, utterance schemas and MWEs, that are combined in different ways. Development has thus been described and analysed as the emergence of new utterance schemas and the combination of such schemas, in an increasing number of ways (and, probably, increasingly fluently) with an increasing number of schematically sanctioned lexical options and intra-turn schematic operations.

Therefore, I believe that the tracing of the emergence of these utterance schemas – the investigation of how they develop – has shed new light on how L2 learning takes place along a usage-based path of routinisation and productivity. I defined routines as concerning both MWEs and utterance schemas, the bulk of portability in L2 learning. Furthermore, I have shown that development is, among other things, a matter of routinisation of these expressions. Development, however, is also learning new utterance schemas and MWEs, and creativity seems to be mostly about using the open

slots in the patterns. Creativity development, then, is using an increasing number of items in the open slots. It is also, as we saw, about combining the schemas and MWEs in new ways.

In terms of UBL, the study has continued where previous ones stopped. In order to investigate if patterns, in time, are linked as constructions (if that is the pivotal process in the ontogenetic development of a L2) this paper has taken concrete pattern instantiations as its starting point and investigated 1) how these emerged into what they are at the time of utterance; and 2) how they are interrelated at the time of the utterance. This investigation has presented, I believe, a window onto how a linguistic inventory might look at a given point in time and how it came to be structured the way it is, taking to heart MacWhinney's (2006) encouragement to look for the roots of the emergent properties of language.

It has been shown how combinability is about putting together chunks rather than lexical items as separate islands. In other words, the lexical items employed are dependent on the patterns known to the language user – and the patterns seem to have been learnt in lexically specific environments as item-based. All in all, these two phenomena – the fact that lexical items sit in certain patterns and the process of learning these patterns as item-based – make for a very context-bound and not very creative linguistic inventory. Spoken language, at least of the kind investigated here, then does not confirm Chomskyan notions of creativity. Language use, and by implication language knowledge (especially to the extent that the two are seen as inseparable in UBL), is much more concrete and stable than is often inferred, which eliminates the validity of the Chomskyan argument from the poverty of the stimulus; language is learnable as an interplay of abstracted utterance schemas, based on exemplar multiword expressions and single lexical items encountered in local environments. Tracing some patterns back to their roots, as it were, we have seen that they invariably stem from very concrete uses and instantiations, supporting N. Ellis' idea that the UBL item-based path of language learning does indeed serve well as a default investigative line of enquiry into questions of L2 development. When and how (or even if) they emerge into more abstract patterns is not easily delineated, the linking of expressions as constructions is still an ambiguous process to go by as default; in stead, it makes more sense to conceptualize emergentism as the linking of singular expressions as utterance schemas, and routinisation as retainment of such singular expressions as (interactional) MWEs. Until we know more about the processes involved here, the best option in

working with language development, L1 or L2, is to take the most concrete starting point possible. If not, important steps in the process towards increased productivity are at the risk of being ignored.

One step that we are now able to take is the one where we can empirically substantiate a justification of ruling out syntactic combinability as the stuff of learning and instead put in item-based utterance schemas as the linguistic material to learn. Pawley and Syder's puzzle of native-like selection, then, is solved by leaving behind once and for all the Chomskyan dualism inherent in the lexicon-syntax division and focus, teachers, learners, users, researchers alike, on a description and analysis of linguistic inventories as item-based as proposed in usage-based linguistics. This is what's truly new!

#### References:

Achard, Michel (2007). Usage-Based Semantics: Meaning and Distribution of Three French 'Breaking' Verbs. In *Collocations and Idioms 1: Papers from the First Nordic Conference on Syntactic Freezes, Joensuu, May 19-20, 2007, Studies in Languages, University of Joensuu*, Vol. 41, Marja Nenonen and Sinikka Niemi (eds.), . Joensuu: Joensuu University Press.

Author, in press. Constructing another Language – Usage-Based Linguistics in Second Language acquisition.

Author, forthc. a: Second language learning as participation and acquisition: towards a new SLA eclecticism.

Author, forthc. b: You no oich – user-based L2 learning: the case of negation.

Author and Colleague. (2007). Are Recurring Multi-Word Expressions really Syntactic Freezes? Second Language Acquisition from the Perspective of Usage-Based Linguistics. In *Collocations and Idioms 1: Papers from the First Nordic Conference on Syntactic Freezes, Joensuu, May 19-20, 2007, Studies in Languages, University of Joensuu*, Vol. 41, Marja Nenonen and Sinikka Niemi (eds.), 86-99. Joensuu: Joensuu University Press.

Bardovi-Harlig, Kathleen (2002). A New Starting Point? Investigating Formulaic Use and Input in Future Expression. *Studies in Second Language Acquisition* 24 (2): 189-198.

Dabrowska, Ewa (2000). From Formula to Schema: The Acquisition of English Questions. *Cognitive Linguistics* 11 (1/2): 83-102.

Dabrowska, Ewa and Elena Lieven (2005). Towards a Lexically Specific Grammar of Children's Question Constructions. *Cognitive Linguistics* 16 (3): 437-474.

- Ellis, Nick C. (2002). Frequency Effects in Language Processing – a Review with Implications for Theories of Implicit and Explicit Language Acquisition. *Studies in Second Language Acquisition* 24 (2): 143-188.
- Ellis, Nick C. and Diane Larsen-Freeman (2006). Language Emergence: Implications for Applied Linguistics – Introduction to the Special Issue. *Applied Linguistics* 27 (4): 558-589.
- Firth, Alan and Johannes Wagner (2007). S/FL Learning as a Social Accomplishment: Elaborations on a ‘Reconceptualized’ SLA. In *Modern Language Journal* 91 (5): 800-819.
- Ford, Cecilia E. (2004): Contingency and Units in Interaction. *Discourse Studies* 6 (1): 27-52.
- Ford, Cecilia E., Barbara A. Fox, and Sandra A. Thompson (1996). Practices in the Construction of Turns: The “TCU” Revisited. *Pragmatics* 6 (3): 427-454.
- Gass, Susan M. (2003). Input and Interaction. In *Handbook in Second Language Acquisition*. Catherine J. Doughty and Michael H. Long (eds.), 224-255. Malden, MA: Blackwell.
- Harder, Peter (2001). The Lexico-Syntactic Symbiosis in a Functional Perspective. *Nordic Journal of Linguistics* 24 (2): 232-242.
- Langacker, Ronald W. (1987). *Foundations of Cognitive Grammar, vol. 1. Theoretical prerequisites*. Stanford: Stanford University Press.
- (2000). A Dynamic Usage-Based Model. In *Usage Based Models of Language*, Michael Barlow and Suzanne Kemmer (eds.): 1-65. Stanford: CSLI.
- Larsen-Freeman, Diane (1997). Chaos/Complexity Science and Second Language Acquisition. In *Applied Linguistics* 18 (2): 141-165.
- (2002). Language Acquisition and Use from a Chaos/Complexity Theory Perspective. In *Language Acquisition and Language Socialization*, Claire Kramsch (ed.), xx-xx London: Continuum
- (2006). The Emergence of Complexity, Fluency, and Accuracy in the Oral and Written Production of Five Chinese Learners of English. *Applied Linguistics* 27 (4): 590-619.
- Lieven, Elena, Heike Behrens, Jennifer Speares, and Michael Tomasello. 2003. Early syntactic creativity: a usage-based approach. *Journal of Child Language* 30: 333-370.
- MacWhinney, Brian (2006). Emergentism: use often and with care. *Applied Linguistics* 27 (4): 590-619.
- Moon, Rosamund 1997. Vocabulary connections: multiword items in English. In *Vocabulary: Description, acquisition, and pedagogy*, Norbert Schmitt & Michael McCarthy (eds.), 40-63. Cambridge: Cambridge University Press.



Nattinger, James R. and Jeanette S. DeCarrico (1992). *Lexical Phrases and Language Teaching*. Oxford: Oxford University Press.

Nesselhauf, Nadja (2005). *Collocations in a Learner Corpus*. Amsterdam: John Benjamins.

Pawley, Andrew and Francis H. Syder (1983). Two Puzzles for Linguistic Theory. In *Language and Communication*, Jack C. Richards and Richard W. Schmidt (eds.), 191-226. Harlow: Longman.

Reder, Steve (2005). The "Lab School". *Focus on Basics* 8(A). URL:  
[http://www.ncsall.net/fileadmin/resources/fob/2005/fob\\_8a.pdf](http://www.ncsall.net/fileadmin/resources/fob/2005/fob_8a.pdf)

Reder, Stephen, Kathryn A. Harris, and Kristen Setzler. (2003). A multimedia adult learner corpus. *TESOL Quarterly* 37 (3): 546-557.

Sacks, Harvey, Emanuel A. Schegloff, and Gail Jefferson (1974). A Simplest Systematic for the Organization of Turn-Taking for Conversation. In *Language* 50: 696-735.

Schegloff, Emanuel A. (1996). Turn organization: one intersection of grammar and interaction. In *Interaction and Grammar*, Elizabeth Ochs, Emanuel A. Schegloff, and Sandra A. Thompson (eds.), 52-133. Cambridge: Cambridge University Press.

Schegloff, Emanuel A., Irene Koshik, Sally Jacoby, and David Olsher (2002). Conversation Analysis and Applied Linguistics. In *Annual Review of Applied Linguistics*, 22: 3-31.

Selting, Margret (2000). The Construction of Units in Conversational Talk. In *Language in Society* 29: 477-517.

- (2005). Syntax and Prosody as Methods for the Construction and Identification of Turn-Constructional Units In Conversation. In *Syntax and Lexis in Conversation: Studies on the Use of Linguistic Resources in Talk-in-Interaction*. Auli Hakulinen and Margret Selting (eds.), 17-44. Philadelphia: John Benjamins.

Taylor, John (2002). *Cognitive Grammar*. Oxford: Oxford University Press

Tomasello, Michael (2000). First Steps toward a Usage-Based Theory of Language Acquisition. *Cognitive Linguistics* 11 (1-2): 61-82.

- (2003). *Constructing a Language*. Cambridge: Cambridge University Press.

Van Lier, Leo. (1996). *Interaction in the Language Curriculum. Awareness, Autonomy and Authenticity*. London: Longman.

- (2002). An Ecological-Semiotic Perspective on Language and Linguistics. *Language Acquisition and Language Socialization*, Claire Kramsch (ed.), 140-164. (London: Continuum)

Appendix: Overview of beginning and end points of recording periods for both focal students:

Recording period 1, Valerio:

**token count:**

MWUs produced by way of

recycled utterance schemas	55	72,4 % (18 containing potential MWEs)
afforded repetitions	11	14,4 %
combinations of previously used...	5	6,6 %
untraceable	5	6,6 %
	76	<u>100,0 %</u>

number of identified utterance schema types: 17

Recording period 6, Valerio:

**token count:**

MWUs by way of

recycled utterance schemas	92	90,2 % (36 containing potential MWEs)
afforded repetitions	1	1,0 %
combinations of previously used...	5	4,9 %
untraceable	4	3,9 %
	<u>102</u>	<u>100,0 %</u>

number of identified utterance schema types: 42

Carlos, recording period 1:

**token count:**

MWUs by way of

recycled utterance schemas	37	86,05 % (8 containing potential MWEs)
afforded repetitions	2	4,65 %
combinations of previously used...	2	4,65 %
untraceable	2	4,65 %

43                    100,00 %

number of identified utterance schema types: 24

Carlos, recording period 4:

**token count:**

MWUs by way of

recycled utterance schemas	108	90 % (24 containing potential MWEs)
afforded repetitions	0	0 %
combinations of previously used...	9	7,5 %
untraceable	3	2,5 %
	<u>120</u>	<u>100,0 %</u>

number of identified utterance schema types: 51

## Chapter 4: ESK3.

Second language learning as participation and acquisition: towards a new SLA eclecticism<sup>33</sup>.

### Abstract:

In this paper, I explore the relationship between local, interactional contingencies and linguistic development in a second language (L2). I will argue that ontogenetic language development is based on numerous occasions of moment-to-moment language use, and, following this observation, I conceptualize the object of research in longitudinal Second Language Acquisition (SLA) research as a hybrid between locally applied usage patterns and application of (the same) usage patterns over time (Author subm.). Studying ontogenetic development, in other words, should not be based on the categories of an a priori formal system; rather, taking the here-and-now world of the speaker as its starting point, this paper advocates the study of L2 linguistic development in terms of an empirically grounded, emergent moment-to-moment grammar, consisting of units of naturally occurring, participant-relevant language use. Drawing on insights from Conversation Analysis (CA) and Usage-Based Linguistics (UBL), this paper investigates the L2 development of an adult classroom learner of English in terms of a process of dual routinization of social activities and linguistic resources. In addition to this dimension of development in interactional competence, UBL's proposed item-based path of language learning, from formulas through low-scope patterns to abstract constructions allows me to properly investigate the nature of the ontogenetic linguistic sediments of the moment-to-moment interactional experience. This paper will ultimately argue for a revision of existing dualisms in applied linguistics by, among other things, making the case for a reconceptualisation of the competence-performance distinction, an abolition of the use-acquisition dichotomy and a notion of learning which makes place for both the learning as participation and learning as acquisition metaphors (Sfard 1998).

### Introduction:

This paper finds itself at a SLA junction. In some respects it follows what has over the last decade transpired as a tendency-turned-tradition of viewing L2 learning as both a social and a cognitive achievement (e.g., Firth and Wagner 1997; Lantolf 2000; Atkinson 2002; Block 2003; Watson-Gegeo 2004; N. Ellis and Larsen-Freeman 2006; Lantolf and Thorne 2006; Larsen-Freeman 2007).

---

<sup>33</sup> Paper under review for publication in *L2 Learning as Social Practice: Conversation-analytic Perspectives* (Working title). Editors Gabriele Pallotti and Johannes Wagner.

In other respects it has fewer parallels; a longitudinal case study of a Mexican-Spanish speaking classroom learner of English, it tracks *both* linguistic and interactional development over a period of almost two years. As such it finds its closest equivalents in similar longitudinal case studies of interactional competence development (Hellermann 2007; Hellermann and Cole *forthc.*). In other respects it is quite its own; combining analytical tools from Conversation Analysis (CA) with Usage-based Linguistics (UBL), a framework for investigating linguistic development in language learning (for L1, see Tomasello 2000, 2003; for L2, see Author and Colleague 2007; Author *subm.*), it tries to capture the relationship between interactional environments and action sequences on the one hand and the individual accumulation of linguistic resources in L2 learning on the other.

Recent years have witnessed an increased interest in the cognitive-social debate in SLA (Larsen-Freeman 2007). A sentiment of a field division is sometimes voiced in the debate, recently in Markee and Kasper (2004) and Zuengler and Miller (2006). For some time, the social viewpoint was seen as a somewhat exotic or peripheral outlook on L2 learning research; it was the challenger's view which now must be reckoned with (Larsen-Freeman 2007). The present paper is introduced into a SLA environment that is thus friendly to experiments along the social-cognitive dimension. The paper picks up the thread of the current debate in two senses: 1) It attempts to heed the call by the likes of Long (1997) and Larsen-Freeman (2007) to empirically validate a socially oriented view of language and language learning; and 2) it picks up Firth and Wagner's (1998) claim that 'functionalist' models of linguistic knowledge, rooted in interactional experience of the learner, are more apt at capturing the dynamics of language learning than a traditional 'structural' one. UBL is precisely such a model, to be delineated below.

First, however, let me briefly consider one key notion which separates the traditional cognitively oriented approaches from that of the 'challenger's', namely the issue of *acquisition*. One challenger position is what has become known as 'CA for SLA' (Markee and Kasper 2004). It is often held against it that the problem in applying CA to SLA lies in its non-cognitive orientation which prevents it from dealing with learning (e.g. Kasper 1997; Larsen-Freeman 2004). While it is true that CA in its original and pure form did not set out to investigate language learning issues, from a CA for SLA perspective this is not a warranted criticism. The concept of learning, however, is typically operationalized by CA practitioners in a way that is quite different from the notion of learning typically found in psycholinguistic research. Abolishing "the computer metaphor" (Block

2003: 97) or the "lone cactus metaphor" (Atkinson 2002: 525) CA for SLA adopt Lave and Wenger's (1991) community of practice framework, focusing on social achievement in a local interactional context in which learning is defined in terms of social actions in social settings and / or increasing participation from peripheral in direction of more full participation in a community of practice (Brouwer & Wagner, 2004; Hellermann, 2006; Hellermann and Cole, *forthc.*; Mondada and Pekarek-Doehler, 2004; Young and Miller, 2004). It is implied in the term *participation* that individuals take part in something that is inherently social in that it consists of more individuals that also *participate*. This participation is, in other words, dependent on what goes on socially among participants in interaction. In this sense, language use, interaction, conversation, can never be situated outside of a local and social context; all (linguistic) actions are locally contingent. The role of the individual is therefore seen as inferior to phenomena that are viewed as socially distributed and co-constructed among and by the participants in interaction. Thus, reflecting its concern with *the collective* whereby its philosophical roots in sociology become evident, this view of learning is arguably exactly where CA for SLA has the most to offer in comparison with more traditional approaches to SLA.

Investigating the development of performance-driven linguistic inventories of the individual L2 learners/users as they collect their biographical anthologies of linguistic experience, as it were, I supplement this view of learning with a more traditional acquisition-based metaphor and maintain that an increasingly fuller participation in interactional settings *results in and is the result of* an emergent portable linguistic ability, something that is available to the individual as he encounters new contexts where he can draw on his former experience. As such, learning is seen as a partly individualized, long-term matter but at the same time one which is dependent on social activity carried out locally. I believe that this is a good way "to show how 'learners' build systematically on experiences from interactions and become more experienced, versatile and competent in the new language" (Brouwer & Wagner, 2004: 42). In such statements I see the potential for gap-bridging across the social-cognitive gap, for viewing individual portability, as it were, as part and parcel of the learning process. This idea of portable sediments of interaction will be explored further in the discussion of data examples in due course.

### The present study.

The data source for the present study is the Multimedia Adult English Learner Corpus (MAELC),<sup>34</sup> which consists of audio-visual recordings of classroom interaction in an English as a Second Language (ESL) context. The classrooms, in which the recordings were made, were equipped with video cameras and students were given wireless microphones on a rotational basis; the teacher also wore a microphone (Reder et al. 2003; Reder 2005). Consisting of recordings from July 2003 through July 2005 this is a longitudinal case study of Valerio<sup>35</sup>, an adult Mexican male learners of English whose L1 is Spanish. The final database of the inquiry consists of transcripts from approx. 30 sessions (each consisting of three hours of recordings, not all of which has been transcribed) in which Valerio is either wearing a microphone or sitting next to someone wearing a microphone. This transcribed database has not been coded for linguistic category information which means that only searches based on exact and concrete sequences of either letters or words are possible.

For the purpose of analyzing and describing the emergent inventory of resources, I invoke UBL. UBL views language knowledge as a structured inventory of symbolic units, whose acquisition is item-based. Often referred to as experience, linguistic knowledge is permeable, flexible, and its development potentially un-ending (Lantolf and Thorne 2006). Therefore, in psycholinguistic terms, it cannot be described as consisting of abstract rules of a finite generative grammar. Rather, linguistic knowledge is psycholinguistically maximalistic; it consists of all kinds of linguistic utterances of varying complexity stored on multiple levels of abstractness, ranging from fixed multi-word expressions to formulaic frames and utterance schemas to fully abstract language knowledge (N. Ellis 2002; Tomasello 2003). However, all L2 learning research can only profitably and empirically justifiably be done in the most specific terms possible; no reference should be made to an abstract level of language knowledge unless it can be empirically substantiated to exist for the linguistic patterns and L2 users under investigation (Author and Colleague 2007; Author subm.). UBL predicts language learning to happen in a slow and piecemeal fashion along a trajectory from formulas (*I dunno*) via partially abstract patterns (*I don't Verb*) to fully abstract constructions (*NP aux neg*) (e.g., Tomasello 2000, N. Ellis 2002).

---

<sup>34</sup> MAELC was compiled and is maintained at The National Labsite for Adult ESOL (known locally as the Lab School). The Lab School was supported, in part, by grant R309B6002 from the Institute for Education Science, U.S. Dept. of Education, to the National Center for the Study of Adult Learning and Literacy (NCSALL) and was a partnership between Portland State University and Portland Community College. I thank Steve Reder and all the staff at the Lab School for granting me access to the data and helping me logistically. This research would not have been possible without their hospitality and assistance.

<sup>35</sup> Valerio is a pseudonym

The data: a window onto the inventory:

Building on previous SLA research which has confirmed the emergent item-based nature of L2 usage patterns according to the UBL path of learning (Author and Colleague 2007; Author subm.), the data for the present investigation concern the item-based nature of the learning of the *auxiliary do*-pattern, not including *aux-neg* pattern. The primary observation in terms of emergentism and item-based learning is that, for the pattern under investigation, there seems to be an initially recurring formula, *What do you say* – a fixed multi word expression (MWE), here operationalized as *a recurring string of words used for a relatively coherent and constant communicative purpose* (Author and Colleague, 2007, Author subm.). The presence of the MWE is striking, representing the first 3 uses of the *aux do* pattern and 38 (54)<sup>36</sup> % of all *aux do* usages in the first period of recording. It is used three times prior to any other *do*-usages. In that sense, it is reasonable to suggest that this is the pattern that sparks off the use/acquisition of more varied and productive *do*-patterns. It is the seed of an emergent pattern. In this initial period it is used 7 times, corresponding to 54 % of *do*-usages in this period.

Table 1 below presents an overview of type-token ratio development for this *aux do*-pattern and, on the right hand side, an impression of the weight of the MWE *what do you say* in relation to the total number of tokens. Tokens are the total number of *aux-do* instantiations whatever the constituents. Types denote the different kinds of instantiations and as such they are more complex; in this table they are distinguished according to both pattern type and main verb. This means that *do you like* and *what do you like* are distinguished, as are *do you like* and *do you say*. A high type-token ratio therefore represents a high degree of productivity across patterns.

	Tokens	Types	Ratio	Number of instantiations of the MWE	Extract presence
Summer 2003	13	5	0,38	7 (ratio: 0,54)	1 – 3c
Autumn 2003	6	5	0,83	2 (ratio: 0,30)	4, 6
Winter 2003-4	6	5	0,83	2 (ratio: 0,30)	5, 9

<sup>36</sup> The statistical uncertainty has to do with the operationalization of MWEs. Instead of viewing the MWE as a purely psycholinguistic phenomenon, stored and retrieved as a whole as opposed to freely combinatorial language production, this section looks at the MWE from a more interactional perspective. One consequence of this is that I chose to conflate the *what* and *how do you say* (which explains the statistical inconsistency) – interactionally they seemed to perform the same operations, so to speak. Furthermore, in terms of sheer numbers, Valerio only used the *how*-variety of the pattern twice (extract 3a).



Spring 2004	25	7	0,28	10 (ratio: 0,40)	7a-c, 8, 10
Summer 2004	18	8	0,44	2 (ratio: 0,11)	11
Summer 2005	15	12	0,80	1 (ratio: 0,07)	12

Table 1: *aux do-usage*. *Extract presence* denotes what instantiations are displayed in what extracts below.

Two parallel tendencies are evident: increasing type-token ratio suggesting increased productivity and reflexive of this, a relative decrease in MWE usage. None of these tendencies are linear, however; fluctuation is the norm as the data confirm the waxing and waning of linguistic patterns as demanded by changing environments (Hopper 1987; Thelen and Bates 2003; Larsen-Freeman 2006; Author subm.). This is especially evident in the type-token ratios which suggest a lower degree of schematicity in late 2004 than late 2003/early 2004. The number of *what do you say* does not account for high number of tokens in Spring and Summer 2004; there must be other recurring expressions in the data in these two periods. This will be returned to below; for now suffice it to say that such unpredictability of linguistic behaviour and development is a core principle found in emergentism (e.g., Hopper 1998; Lantolf 2005), Chaos/Complexity Theory (Larsen-Freeman 1997), and Dynamic Systems Theory (Thelen and Bates 2003; Larsen-Freeman 2006; De Bot et al. 2007).

As shown in the column on the right, the MWE displays a slightly descending presence in the data from 54% of all *do* usage initially to 40% roughly one year into the recording period – until it abruptly plunges to the brink of disappearance. This finding is reminiscent of a process which also seemed to hold sway in an analysis of *can*-patterns in Carlos, another Mexican classroom learner of L2 English (Author subm.). Thus, seeing it again here confirmed the speculation from that study on the *can*-patterns, that language knowledge, also for L2 learners, is constantly in flux and under revision as changing environments call upon different linguistic resources (Hopper 1998). Linguistic patterns go in and out of experience; an idea which is at the core of the emergentist assumption that interaction is a constant source of renewal for the individual linguistic inventory, this goes well with a locally contextualized notion of language knowledge in which linguistic expressions are seen as fundamentally tied to specific situations; i.e. interactionally contingent.

Furthermore, the manner in which the larger patterns emerge (here *aux do*) stands in contrast to common findings in research into the role of MWEs (usually referred to as Formulaic Sequences (Wray 2002; Schmitt and Carter 2004)) in SLA. This research commonly views formulas as being

beyond the current interlanguage capabilities of the learner (e.g. Myles et al. 1998; Bardovi-Harlig 2002; also discussion in Author subm.) The present data, on the other hand, show how the MWE (here *what do you say*) is itself an integral part of the developing linguistic inventory; it is not somehow beyond a current combinatorial interlanguage system.

Table 2 (appendix) shows in more detail this emergent nature of *aux do*-pattern development. Numbers in red represent use of any of the five pattern types in use since the first recording period. From Table 2 it is evident that the most advanced uses are traceable to previous experience; the linguistic inventory is constructed in this stepping-stone fashion as the emergent individual grammar is called upon in a variety of usage events (Langacker 2000; Author subm.). Table 2 also displays the item-based nature of the learning trajectory from the MWE toward a richer inventory of interrelated linguistic expressions and patterns, perhaps increasingly schematic structures. It lies beyond the scope of the present paper to go into a detailed discussion concerning the degree of schematicity of the underlying language knowledge in the inventory, which is a vexing issue (Lieven et al. 2003; Author subm.). For this particular pattern it seems that *do you Verb* is the kernel, *do you* perhaps a totally fixed part, with the possible addition of *wh*-question markers of various kinds. Thus, we might speculate that for Valerio the central *aux do*-pattern is an utterance schema (Tomasello 2000; Author subm.), catalogued along the following lines in the linguistic inventory: (*WH*) *do you Verb?*

*Aux-do* development, then, item-based in nature, seems to hinge on an initially highly recurring MWE<sup>37</sup>. Table 2 reveals the further existence of other potential MWE candidates, namely *what do you write*, *do you like*, *do you have*, which are temporally unstable; *what do you write*, *what / when do you use* recurred in Spring 2004, and *do you have* is the recurring expression making for the high number of tokens in Summer 2004. This instability, or fluctuating nature of the MWEs, was also an issue in Author (subm.), where it was found that Carlos's (another classroom student) *can*-pattern emergence was traceable to a few initially recurring MWEs – all of which eventually disappeared from the data. The explanation for this fluctuation was found in the recurrent classroom activities in which the MWEs were deployed. When such activities were not on the agenda, the MWEs were seemingly discharged. This implies that MWEs must be defined interactionally as well as with

---

<sup>37</sup> In fact, there might be one more MWE, namely *do you like np/v/ø*; so there may be two items from which the emergence of the pattern originates. The initial existence of two exemplar patterns does not change the view of development as item-based, however.

reference to issues of storage and processing. In other words, at those points in time where the MWE is in use we cannot convincingly argue that it is not stored as a unit – especially since it is used across some time and in multiple contexts (which will be shown in data extracts below). This storage as whole may be exactly what makes it so useful interactionally, yet again underlining the reflexive nature of ontogenetic language development and interaction, of individuality and sociality. Thus, the reasons for the temporal instability of the patterns are probably not to be found in psycholinguistic terms of storage and retrieval; rather, they are more likely to be a reflection of the focal student's moment-to-moment communicative needs in the classroom. This, however, is a statement that requires empirical substantiation, and in Author (subm.) the claim was justified by showing a co-occurrence of given MWEs and recurring classroom activities, thus supporting an argument in favour of a more locally contextualized notion of L2 development. In this, the study confirmed an on-going study which established correlations between routinised activities and routinised expressions (Author and Colleague2 2007). Parallel to the findings concerning interactional routines in much language socialization research (e.g. Kanagy 1999; see Zuengler and Cole 2005 for an overview), this common tendency in these various studies sowed the seed of the present investigation.

The present paper, then, takes the implications a step further to see how far we can take the endeavour of co-materializing aspects of interaction and use and aspects of acquisition and knowledge, long held to be antagonistic notions in SLA, to trace the co-emergence of interactional contingencies and individual linguistic inventories. It is inferred here that proponents of both social and cognitive approaches to the study of L2 learning could benefit greatly from indulging in this particular problem. The present paper will make the case in favour of finding a place to study L2 development in all forms, namely the place where interactional routines and linguistic productivity meet. Here is where we might empirically substantiate the non-division of use and acquisition, competence and performance, combine the metaphors of learning as participation and learning as acquisition – if we can get at the nature of interactional contingencies and individual linguistic inventories as they co-develop.

#### Zooming in on the inventory in use:

One way to get at these interactional contingencies is to put the data under the scrutiny of a conversation analysis (CA) inspired analysis to investigate issues pertaining to sequential

placement. In the interest of clarity it should be stressed that this investigation is not strictly CA; rather, some concepts often used in micro-analysis have been borrowed for the present purposes. The common-ground shared by the present research and proponents of micro-analysis for L2 studies, Firth and Wagner (2007), is the conceptualisation of language learning as emergent, constant, and never-ending, which justifies this paper's curiosity to see whether, quoting Wagner and Gardner (2004: 14), "a micro-analysis of second language conversations can enhance our understanding of what it means to talk in another language, by broadening the focus beyond the sounds, structures and meanings of language to encompass action sequences, timing and interactivity (...)". In other words, the CA insight that conversation is organised in action sequences, and the ensuing discovery that what people accomplish through language depends on sequential positioning of linguistic items (Sacks and Schegloff 1979, quoted in Ten Have 1999), is what makes micro-analysis relevant for the study of L2 interactional data.

The micro-analytic terminology applied here includes sequences, openings and closings, and participants' orientations. Conversations consist of sequences, which consist of the participants' turn-taking (Sacks et al. 1974). The sequences of conversations have most famously been described in terms of phone conversations which, to varying degree and among other things, consists of sequences such as 'summons-answer sequence', 'greetings and how are you sequence', 'the reason for the call sequence', and a 'closing sequence'. Sequences, then, are not ultimately defined by topic (Schegloff 2007) but by the action that is carried out and accomplished in the interaction. It is in this sense that the notion of sequence is used here – a series of turns-at-talk in which some action is carried out and accomplished by the speakers-hearers.

So, in the extracts below the target utterance is the MWE *what do you say*, the focus being on what it does and where it does it in terms of its sequential position and the orientations of the participants. The interest, then, lies in the interplay among the MWE, sequential organization, and social activity, but NOT in social order per se. On a more epistemological note, the current interest is also in the interplay between the social nature of the classroom interactions and the nature of the contributions of the individual participants in interaction, the underlying assumption being that these co-develop in an equal partnership, as it were, in which none is given prominence over the other.

Extract 1 below is from Valerio's very first day in ESL class. In the preceding discourse, a task in which participants talk about favourite holidays / days of the year, Valerio has been 'interviewing' Ang, who does not immediately begin to interview him back. Instead Valerio now self-selects as the next speaker (l. 1) to tell Ang about his favourite day of the year, his wedding anniversary.

Extract 1 (July 1 2003)<sup>38</sup>:

- 1 Valerio ((writes)) <spn> aih ((slaps himself on mouth)) uh hh:h for me: December (1) eight  
2 Ang ((writes)) December ((writes)) December?  
3 Valerio eight (3) because (1) is (1) anniversary (3)  
4 Ang uhuh (2)  
5 Valerio for the marri<sup>↑</sup>ed (2)  
6 Ang anniversary?  
7 Valerio for the marri<sup>↑</sup>ed  
8 Ang ((waving hand)) for is *para*  
9 Valerio for ((frowns)) (2) um:  
10 (5)  
11 Ang xxx ((turns around to her own desk, writes))  
12 Valer → agh: ((leans over to see what Ang writes)) (4) teacher what do you say for  
13 (2)  
14 Ang anniversary[: and marry ((pointing at Valerio))  
15 Valerio [for anniversary married?  
16 Dom it's it's anniversary.  
17 Valerio anni[versary?  
18 Dom [wedding anniversary  
19 Ang wedding? ((picks up paper))  
20 Valerio wed[ding?  
21 Dom [wedding (+) so let [me write it down for you]  
22 Valerio [wedding (.) ah yeah yeah]  
23 Dom wedding [anniversary  
24 Valerio [wedding is here ((shuffles paper, writes)) no (2) uhuh ((points to board,

<sup>38</sup> Transcript conventions: () = pause; (...): transcriber's comments; (...): uncertain transcription; : = prolonged phoneme; underlined item(s): stressed item(s); [ ] = overlap; arrows up / down = rising / markedly falling intonation; → = target utterance; \*...\* = silent voice; U = unidentified student. All names are pseudonyms.

25            looks at Ang)) wedding anniversary (1) is wedding anniversary  
26 Ang        ((writes))

In line 3 Valerio uses the term *anniversary* which, after a 3 second pause in which Ang writes down Valerio's answer, is acknowledged by Ang (l. 4). Then troubles arise as Valerio (l. 5) goes on to add *anniversary* to his previous turn *for the married* (he pronounces the final syllable [id]) to denote that he is talking about his wedding anniversary. Ang initiates a repair sequence, *anniversary?* in l. 6, and in return Valerio offers a repetition of *for the married?*. At this stage it seems that the participants are negotiating how to express 'wedding anniversary' properly in English. The immediate focus of Ang's repair at this stage seems to be Valerio's use of *for* rather than *married*, as she gives him the Spanish equivalent (l. 8.). Valerio frowns, repeating *for*, and Ang turns around to her own desk (at the same time facing the whiteboard). Eventually, the teacher approaches the pair, and the repair sequence results in a side sequence (Brouwer 2004) in which a 'lexical inquiry' is carried out as Valerio summons the teacher, explicitly inviting for help (Brouwer 2003) to get at the target word, 'wedding anniversary' (l. 12). Valerio, in other words, performs a self-initiated other-oriented, post-trouble problem-solving activity by using the MWE *what do you say (for)*. The two students go on to co-construct the lexical inquiry, indexing it as a joint problem, (ll. 14-15) resulting in the teacher giving the English term (ll. 16-18). The sequence is eventually closed down successfully as the students both seem to accept that the trouble item is in place (ll. 24-26).

In the next interaction (extract 2), occurring 17 days after the first one, Valerio uses the MWE twice for a related purpose, first with his partner in the task, next with another classmate, in a brief side sequence requiring a summons (as was the case in the previous interaction in which Valerio summoned the teacher).

Extract 2 (July 18 2003):

1 Ang        mhm (1) okay what did you\_ you do xxx  
2 Valerio    what did you xxx  
3 Ang        mhm  
4            (2)  
5 Valerio    I: visit dents (2) I am visit des dents  
6 Ang        dens?

7 Valerio dens dent dent

8 Ang ((writing)) dens dens

9 Valerio dents uhuh

10 (1)

11 Ang dens

12 Valerio dents

13 Ang <spn = *asi no se escribe?*> dens? ("this is not how you write it, is it?")

14 Valerio <spn *denti:sta*>

15 (.)

16 Ang <spn *ah es con teh*> ("oh it's with tee")

17 Valerio uhuh

18 Ang dents

19 Valerio a:nd (5) <spn *bueno*> yesterday <spn *ya*>. yes. he visit dents. (2) for evening?

20 (1)

21 Ang <spn (*muchas cosas hicistes?*)> ("(many things) you have done?")

22 Valer mhm <spn *estuve en la casa de xxx*>

23 Ang xxx

24 Valeri → <spn *ah entonces es*> I have (1) I am ho: (2) I hoh what do you say. ha:ve I'meh I

25 bdrbdrbdr (1) in my home (3) the morning in my home and visit dents ((points to

26 Ang's paper))

27 Ang <spn *no es*> [next

28 Valerio [visit dents

29 Ang next <spn *del dentista*>

30 Valerio no next <spn *fue visitar al dentista*> ("was a visit to the dentist")

31 Ang <spn *y luego*> (and then)

32 Valerio I visit ((writes))

33 Ang ((writes)) e:hm (+) in the morning (3) <spn *es que xxx visit e: correcto?*>

34 Valerio ((grabs and looks in electronic dictionary)) to be? (1) <spn *qué es el verbo to be? estar*>

35 Ang uh:h

36 ((both look in electronic dictionary))

37 Valerio <spn *estabamos no da*> ([we were] doesn't work) hm hm (17 ((looks in ED))) Lore

38 → ((establishes eye contact with Lorena who sits across the room)) what do you say

- 39            <estuvo>  
 40 Lore      <estuvo>? (.) I went?  
 41 Valerio    I went?  
 42 Lore      uhuh  
 43 Valerio    alright thank you () <spn ya más fácil>  
 44 Lore      <spn si si sabe cómo se escribe?>  
 45 Valerio    I went ((nods))

In the task the students are instructed to talk about what they did the day before. Valerio tells Angelina that he has been to the dentist. This creates some trouble as the students jointly agree on its pronunciation and spelling (ll. 1-18). This sequence is emphatically closed by Valerio in l. 19, as he, partly in Spanish, says that 'yesterday' is a wrap – he visited the dentist. The 3<sup>rd</sup> person usage is task-specific; the students are supposed to collect information from each other so the 3<sup>rd</sup> person probably comes from Ang's writing.

Valerio then goes on to talk about what he did in the evening. This Angelina responds to by giving a positive assessment (l. 21). In the following turn Valerio says in Spanish that he was at somebody's house (inaudible name), which Ang responds to in another inaudible turn. This is followed by Valerio refocusing on the task, also partly in Spanish, *entonces es I have*. Next, Valerio starts hesitating, marked by the stretching of the vowel sound in *I am ho*: which eventually turns into a word search, in the midst of which we have the target expression *what do you say?* pronounced with slightly falling intonation. As opposed to the previous interaction, however, this time the MWE is not used as an invitation for help (Brouwer 2003), and it is not oriented to as such by the co-participant. Valerio here makes it public that he is "doing thinking" (Houtkoop-Stenstra 1994, quoted in Brouwer 2003: 538) while keeping the floor. This is achieved by the vowel stretching, a common pre-indicator of a word search (Schegloff 1979, quoted in Brouwer 2003: 541) and by using the MWE but also by other modalities, such as falling intonation and, especially, by avoiding eye contact. In the previous turn exchange, in Spanish, there was eye contact between the participants; however, as Valerio starts searching for the word his gaze wanders off. Accordingly, Ang does not intervene in Valerio's word search and it seems he does not expect her to. The eye contact is reestablished later in Valerio's turn, during the 3 second pause, at which stage the focus is no longer on the word search but back on the task itself. They then seem to be summarizing what information Ang has retrieved from Valerio (ll. 27-32), and then Ang, l. 33,



initiates a new search for a lexical item as she seems to express doubts that *visit* is the right verb; this turn, unfortunately, is partially inaudible. This eventually results in Valerio employing the MWE in a manner similar to Extract 1 above where he first summons the intended recipient before asking for the specific item (l. 37). The summons, which receives a non-verbal answer in that a mutual gaze is established, and the MWE *What do you say* (l. 38) open the side sequence which is eventually closed by Valerio appreciating the help expressed by *alright thank you* and the nodding (l. 45) in response to Lore's 'reopening' question *do you know how to write it?* in Spanish.

At this stage, then, we have three instances of the MWE, two of which recur in comparable sequences where Valerio explicitly invites for help to do a lexical inquiry. The third instance also takes place in a lexical search environment, but it is carried out in a manner which is not invitational as Valerio uses it to display that he is 'doing thinking' .

Extract 3a, Aug12 2003:

- 1 Dom here is bread  
2 Valerio hm  
3 Dom here is some bread ((showing them book))  
4 (1)  
5 Valerio for  
6 Oliv xxx ((looks in teacher's book, nods))  
7 Dom mhm  
8 Valerio teacher  
9 Dom mhm  
10 Valeri → here is bread all debr all de:hm (1) hh in Mexico is different names in: ((points to  
11 teacher's book)) the brea:d or for the: ((looks in partner's direction)) nhah ((slight  
12 laughter)) how do you say hh ((looks down)) how do you say <spn *panaderia*> (2) for  
13 the make () bread?  
14 Dom mhm [a bakery]  
15 Valerio [is ] is de is xxx an[:d  
16 Dom [mmm ts  
17 Valerio it's very different↓  
18 Dom yeah there are many [different breads]

19 Lorena [teacher ] I have a question

20 Dom yes ((leaves V's table))

In this extract, 1 month and 12 days after extract 1, the MWE is used with *how* rather than *what* which makes for a more native-like pattern. However, this pattern has not substituted the older pattern which is reused in later interactions, to be returned to. This 'how'-pattern, on the other hand, is not found again in Valerio's data. Leading up to this extract, the teacher has been assisting Olivia with a troubling word, *bread*. This item, part of a sentence which the students are asked to complete in the task, did not seem to present any problems to Valerio. As Olivia acknowledges the teacher's help (l. 6), Valerio summons the teacher and starts talking about "different breads". At this stage there is an established interactional space in which the two students and the teacher co-participate; Ang and the teacher are both potential recipients of Valerio's turn, as implied by their gazes. The first trouble indicator marked by a stretched vowel comes in *all de:hm*, at which point Olivia leaves the established interactional space, turning her gaze downwards. After a 1 second pause, Valerio restarts his turn *hh in Mexico is...* at which point Olivia realigns with Valerio as she turns her gaze towards him; at the mentioning of *Mexico* she is again a ratified co-participant (Goffman 1981), sharing membership knowledge with Valerio (both are from Mexico). Following the next trouble indicator *the:*, and in this respect this extract is similar to the previous one, the MWE, used twice, sits mid-turn as a display of doing thinking. Valerio's head-turning towards Olivia may be a search for support but the fast transition via in-breath into the 2<sup>nd</sup> instantiation of the MWE suggests that he is here signaling 'more to come' (Schegloff 1996) as he is doing a step-wise withdrawal from the interactional space which he has shared with the teacher since the summons. In that sense, given the lack of reaction from Olivia and the fast transition into the 2<sup>nd</sup> instantiation of the MWE, the shifting gaze suggests that he is momentarily excluding the teacher from reciprocity. She is reestablished as potential recipient as Valerio looks up immediately following the second MWE and co-occurring with the Spanish *panaderia*. The teacher then gives her solution candidate after Valerio has elaborated on what he is after.

Arguably, the reason why Valerio repeats the MWE combined with a self-oriented physical presence in this interaction is because he intends to keep the word search non-invitational. When he is more other-oriented, he may use, together with mutual gazing or other forms of bodily outward orientation, a summons before involving other participants in his word search. In the next extract,

recorded later the same day, the MWE has a flat, perhaps slightly falling intonation, but his physical presence alone seems to be enough to signal an invitation to the participant. Interestingly, then, at this stage, we seem to have a MWE, perhaps two MWEs, accompanied by different gazes each achieving different interactional functions.

Extract 3b (Aug 12 2003):

- 1 Valerio in this country ((reading in a whisper, leaned back))
- 2 Olivia ((writes))
- 3 Valeri → I no [understand hh what do you say ((sits up straight))]
- 4 Olivia [((increasingly orienting to Valerio)) ] ((mutual gaze))
- 5 Valerio hh in this country ((looks down at paper, starts reading aloud))
- 6 Olivia ((looks down at paper))
- 7 Valerio you can orden some () keends↑
- 8 Olivia keends
- 9 Valerio of food () by telephone boot but
- 10 Olivia ((reading)) but it's better cook in your house () <spn *es mejor cocinar en tu casa*>
- 11 humhumhum ((laughter))
- 12 (3) ((Valerio moving his shoulders))
- 13 Valerio hm hm [((nods, taps paper with pen))]
- 14 Olivia [order some xxx
- 15 Valerio hh order [some kind order]
- 16 Olivia [xxx ]
- 17 Valerio order [some keends ]
- 18 Olivia [((leans towards Valerio)) yes↑]
- 19 Valerio order some keends what do you. order some keends
- 20 Olivia order [some
- 21 Valerio [*bueno* order yes some [keends
- 22 Olivia [some some keends <spn *no atrapé xxx*> (1) some
- 23 keends↓ (+) keends? some keends <spn *xxx*> hehm
- 24 Valerio ((whispers)) (3) ((shows dictionary))
- 25 Olivia <spn>
- 26 Valerio a:.....:hh (1) bo[o:...

27 Olivia [xxx ((writes, looks at Valerio's paper, puts down pencil))  
28 Valer → boo: alright ((erases, hits table with side of hand 4 times, writes *I don't like because it*  
29 *has different taste*))

In extract 3b, the target expression is in line 3, following private reading and followed by reading aloud. Previously in the interaction there has been some task-solving and a pause and now Valerio opens a new sequence. At first, leaning back, he reads in a whisper (*in this country*). Then he starts to sit up straight saying *I no understand hh what do you say* and then, marked by a pitch reset, starting to read out loud. Olivia reacts to *I no understand* and Valerio's changing posture by increasingly orienting to him and at the end of *what do you say*, a mutual interactional space has been established. Valerio has trouble understanding the sentence which the task instruction tells the students to 'end logically' by adding a *but*-clause to it. This much also seems clear to the students who co-construct the task activity as evidenced in l. 10 where Olivia finishes Valerio's turn by offering her sensible ending to the sentence. After 3 seconds of pondering over Olivia's solution (marked by the shoulder dance), Valerio accepts it (acknowledgment token in l. 13).

Still, Olivia picks up the lexical inquiry thread instigated by Valerio, whose invitation to do a collaborative lexical search was initiated by the MWE accompanied by an active body presence. Identifying the trouble-source as 'keends' is achieved in a co-constructed manner, first in ll. 7-9, and then again in ll. 14-18, as Valerio and Olivia increasingly orient to this specific item. Valerio makes it very specific in l. 21 where he expresses that *order* is not the problem (in much the same as way as he closed the 'yesterday-sequence' in a previous interaction) and in the following turn it seems that Olivia aligns with this as she expresses, in Spanish, "not having caught" something (partially inaudible). The sequence is finally closed as Valerio completes the sentence, writing *but I don't like because it has different taste*. It is thus not entirely clear what he makes of the original sentence, as his solution is not entirely 'logical' in the sense the task instruction was to end the sentence in a logical manner. Arguably, Olivia's solution was more apt – and in a later sequence accepted by the teacher (who, unfortunately for the present purposes, does not check Valerio's solution). Even though the lexical search prompted by the MWE and co-constructed by our participants is not evidently solved, which is not uncommon (Goodwin and Goodwin 1986, quoted in Brouwer 2003: 541), the role played by the MWE as inviting for help in specific lexical search environments remains fairly well-established.

Extract 3c (Aug 12 2003)

- 1 Olivia Valerio do you like to go to the shopping?  
2 (2)  
3 Valerio ye[:s I do hrhehhehheh]  
4 Olivia [hmmhmhehehheh ] yes what kind of shopping? () xxx () what kind  
5 Valerio uhm for closeh  
6 Olivia xxx  
7 Valerio [closeh?  
8 Olivia [closeh? ((shrugging gesture))  
9 Valerio always closeh?  
10 Olivia ((shrugs)) hn[:  
11 Valerio [o:h no [always ((shaking head)). is ((points to Olivia, waving fingers))]  
12 Olivia [all ((fiddles with shirt)) <spn *come se dice ropa*> (2) ]  
13 clothes?  
13 Valerio cl[othes  
14 Olivia [clothes ah:m ((nods))  
15 Valer → clothes closeh clothes (1) uhm ((waves toward Olivia)) what do you say no always  
16 closeh  
17 Olivia hm:  
18 Valerio ehm [no is always  
19 Olivia [((shrugs)) <spn *no sé hm*> mhm ((nods))  
20 ((both orient away))

In this extract, still on the same day, Ang asks Valerio about shopping. Valerio tells her that he likes shopping for clothes. The initial part of the interaction concerns agreeing on the word clothes. Then, l. 9 marks an incipient interactional split, as it were, as Valerio seems increasingly concerned with *always*, which is evident in l. 11 where the pitch reset between *o:h no always* and *is*, accompanied by the waving of the fingers in Olivia's direction, gives away his focus on *always*. In overlap, Olivia continues the *clothes* discourse, perhaps because she was not entirely happy with Valerio's pronunciation; in any case, her pronunciation is more native-like than Valerio's. Having repeated Olivia's *clothes*, and after collaboratively ending the sequence (ll. 14-15), Valerio then opens a new

sequence, *uhm* followed by the wave toward Olivia and the MWE, performing yet again the invitation for help. It is not immediately evident from this extract, but Valerio is probably after the word 'kinds' – which previously caused them problems, as already displayed in extract 3b. This hypothesis is further supported by an interaction, approx. 6 minutes earlier, where the teacher asked Valerio what he likes. When he answers 'chocolate' she asks him what kinds. He says 'always chocolate' and she corrects him, 'all kinds of chocolates', which he then repeats, and the teacher moves on. So in this interaction, Valerio probably remembers that it is not the word 'always' he is after but the words 'all kinds' – as in 'all kinds of clothes'. Even though this is not clear to Olivia either, the function of the MWE remains well-established and the sequence is closed down even if Valerio does not get what he is after.

Extract 4 (Nov 18 2003)

- 1 Lore ((points at Valerio's partner)) she wants to speak spanish  
 2 Partner ahh heh ((laughter))  
 3 Valerio → ((nods)) yeah ((waves at Lore)) what do you say muy pronto () in english  
 4 Lore muy pronto? fast () very fast  
 5 Valerio very fast ((private speech)) very fast ((whispering, to partner))  
 6 Partner what  
 7 Valerio ((whispering)) very fast () muy pronto ((pointing to partner's paper)) () very fast this  
 8 one () no no this one ((pointing to partner's paper))  
 9 Partner no I know I know I know

In extract 4, 4½ months after first occurrence, the target MWE is still used to ask for a lexical item but, unlike previous extracts it happens in an off-task environment. As such, the bound activity which it initiates is driven more by Valerio's own curiosity than some desire to solve a task. The situation still contains a repairable (e.g., Schegloff 2007), however, and the phrase is used to perform a self-initiated other-repair. Valerio's partner has expressed a wish to learn Spanish and Valerio is keen on giving her a translation of a Spanish phrase in her papers. However, Valerio first needs to get the translation from a Spanish speaking peer, which he does without pausing or otherwise marking that a word search activity is going to be carried out. Rather, the expression sits inside the ongoing interaction, as it were, reflecting Valerio's ability to carry the expression into a

new kind of environment. The wave (l. 3) is the expressive gesture signaling the invitation for help to get at the lexical item.

Extract 5 Mar 9 04

- 1 Valer do you like meh eh food Mexican?  
2 (2)  
3 David maybe but I didn't uh but I do not taste  
4 Valer ((nods))  
5 David maybe  
6 Valer → you yo↓u:: ts () what do you say ((looks down)) (1) you eating mex you eating food  
7 Mexican?  
8 David no ((shakes head)) never  
9 Valer never?  
10 David yes  
11 Valer I like I you like eh food Chinese?  
12 David Chinese food yes  
13 Valer ((writes))

This extract, recorded eight months and nine days after extract 1, shows that the MWE as a tool to perform a self-initiated self-repair in the form of a self-oriented lexical search is retained. The sequential position, after a first trouble indicator and followed or accompanied by some kind of inward physical presence as well as the co-participant's orientation to the MWE are all parallel with the previously seen interactions in which the MWE was used in the same way. In this respect, there seems to be a correlation between interactional circumstances and the relevance of the MWE. David's *yes* in l. 10, it should be noted, is task-specific; it is an affirmation that *never* is the correct piece of information for Valerio to write down about David in this task.

So far we have seen, then, that the MWE is linked to certain sequential environments in which it performs certain activities and from which it derives its functions – these things hang together and they are what the participants primarily seem to be orienting to. The data have shown that the MWE may be used to ask for help (teacher / peer) or to display "doing thinking". The explicit invitation for help is always pre-indexed by some kind of interactional work, either a summons, a gesture, or a





9 Ian can you lend me  
 10 Valerio n[o: ((shakes head; points to Ian's book))  
 11 Ian [can I [borrow]  
 12 Valerio [borrow]  
 13 Ian I borrow you lend me () is sa[me  
 14 Valerio [I lend?  
 15 Ian yeah you lend you lend is get is [get somebody  
 16 Valerio [alright ((looks in book))  
 17 Ian [I borrow I I ((motions receiving))  
 18 Valerio [ah yeah yeah yeah yeah yeah  
 19 Ian I need to get something  
 20 Valerio yeah I understand xxx example  
 21 Ian uhuh  
 22 Valerio thank you  
 23 Ian you're welcome  
 24 Valerio I confuse  
 25 Ian hehheh me too

In this extract, Valerio is aiming at something that has passed in the previous interaction, as he tries to get Ian to say something which he said earlier. The task revolves around the students making requests containing *can* and *borrow*. Earlier in the interaction, Valerio asked Ian about *eggs* and Ian made a joke. In this interaction, Valerio deploys the MWE not only to ask Ian to repeat what he said earlier (l. 2), he seems to challenge (Schegloff 1995; Koshik 2003) the correctness or relevance of a previous utterance made by Ian. This other-initiated self-repair concerns something that was said quite a few turns earlier, and is thus reminiscent of what Schegloff (1992) called *third position repair*, and even though Ian seemingly produces a relevant proposal for a solution, i.e., he orients to the repair initiation as such, the challenge-aspect of the interaction becomes increasingly clear as Valerio openly performs an other-initiated other-repair (l. 6) and the interaction becomes an exchange of other-repairs as the two participants collaboratively turn it into a more principled discussion of the difference between *lend* and *borrow*. The discussion finally closes as Valerio acknowledges that Ian is right and thanks him for his help.

Approximately 6 months later, the MWE is used several times in similar environments, displayed in extract 7a-c below. All recorded on the same day, they show how Valerio's use of the MWE as an other-oriented repair initiator is becoming entrenched.

Extract 7a May 4-04 ca. 2400

- 1 Sam no no just talking about something [you learned outside of this ((pointing down))  
2 Valerio [for this week? () for this week?  
3 Sam xxx ((shakes head))  
4 Valerio maybe I: (5) I don't know heh heh he[h hh  
5 Sam [you don't remember?  
6 (2)  
7 Valerio → mm (3) you you: ((leans forward, starts looking at Sam's paper)) what do you say for  
8 this week  
9 Sam out of English class I learn examples that uh I wrote here ((points to paper))  
10 Valerio I: () no but ((shaking head, puts hand on Sam's paper)) you write what are you doing  
11 for the five weeks before ((pointing to paper))  
12 Sam what's that  
13 Valerio you write what are you doing for the five five weeks before no  
14 Sam ((pointing to paper)) I () read this read this  
15 Valerio I was learning this word out of this class xxx ((reads Sam's list of learned words))

In extract 7a, the students are talking about what they have learned recently. The teacher has instructed them to talk about what they have learned over the past five weeks as well as what they have learned outside of class the previous weekend. Sam asks Valerio what he has learned. At first doing a repair (l. 2), he seems, for some reason, unable or unwilling to give an answer (l. 4). He then seems to turn to Sam's list of recently learned words, displayed his change in posture (l. 7). The MWE in l. 7-8 is used to ask what Sam learnt this week, Valerio apparently assuming that Sam's list of words covers the previous five weeks. While it may not be entirely clear what the MWE is supposed to do, Sam starts orienting to his paper immediately following the MWE, responds to the MWE as a repair and offers essentially the same candidate solution in l. 9 and l. 14, namely to direct Valerio's attention to the words he has written down as learned. Seemingly satisfied, Valerio, in the interaction which follows this extract, starts asking Sam about the specific words.

Extract 7b May 4-04 ca. 4400

- 1 Sam we have to use it for for ((points to book))  
2 (1)  
3 Valerio for?  
4 Sam I I got ((raises hand; orients out into the classroom))  
5 Valerio → what did you say use it for for  
6 Sam what's that?  
7 Valerio → what do you say for for? You no remember what do you say for [for?  
8 Sam [yeah I have been  
9 living in Portland for for [xxx]

Extract 7c May 04 04 ca 5130

- 1 Valerio but you rep repeat the question I have learned no how long how long  
2 Sam how [long  
3 Valeri → [a:h yeah yeah yeah yeah yeah yeah yeah you coop you coop what do  
4 you say?  
5 Sam you coop?  
6 Valerio no  
7 Sam copy?  
8 Valerio no no no no no  
9 Sam xxx  
10 Valerio you repeat  
11 Sam what question?  
12 Valerio you repeat before  
13 Sam how long I've been [xxx  
14 Valerio [no no no no () no for your question () you repeat mm xxx I am  
15 tired hh  
16 Sam copy copy?  
17 Valerio no:.....  
18 Sam what do you want to say?  
19 Valerio that's alright ((withdraws physically))

In the above extracts the students are working on a task involving the production of sentences in present perfect / present perfect continuous. In Extract 7b, they have been discussing whether to use *since* or *for* in their constructed examples. Valerio asks Sam if he can use the present perfect continuous with *since* when Sam tells him to use *for* (l. 1). Valerio reacts to this by initiating a repair-sequence which Sam, even though he repairs Valerio's repair (l. 6), eventually completes (ll. 8-9). In Extract 7c, the discussion on how to construct questions with the present perfect concerns whether to use *how long* in all the examples. As in Extract 6, Valerio uses the MWE to get at something previously said by the co-participant. Whereas the trouble was solved in Extract 6, this does not appear to be the case in this 8c. At first Sam offers two solutions (l. 5 + l. 7), tries to get more information (l. 11), and offers solutions yet again in l. 13 and l. 16. None of the solution candidates are acknowledged by Valerio, who eventually closes the sequence without having accomplished a satisfying outcome (l. 19). The MWE itself, however, works seemingly unproblematically as an other-initiated self-repair – even though the trouble is not solved.

Then, 10 days later, the following interaction takes place:

Extract 8: May 14 04

- 1 Sam            what do you think xxx person (5) what do you think?
- 2 Valerio      ((looks away, writes)) sowld () showld
- 3 Sam            ((orienting to paper)) okay you should listening to old generation
- 4                (1)
- 5 Valerio → what did you say?
- 6 Sam            you know it has these are these are le learning mistakes these are not right ((pointing
- 7                to paper))
- 8 Valerio      this is no right?
- 9 Sam            no it's right ((shakes head))
- 10 Valer → hm () what do you say
- 11 Sam          what () [xxx
- 12 Valer →            [what do you say
- 13 Sam          I just read these sentence ((pointing to paper))
- 14 Valer → and what do you say?
- 15 Sam          hh I s I I think this would be listen () you should listen to old genera[tion

16 Valerio

[yeah () you

17 should listen old generation or to old generation

The first use in this interaction is parallel to the two previous ones; it sits in a self-initiated other-repair sequence. The students have been instructed to correct four sentences which the teacher has retrieved from the students' writing samples. Sam reads aloud from the task paper (l. 3). After a 1 second pause, indicating trouble, Valerio initiates the repair (l. 5), using the past tense in much the same way as in Extract 7b<sup>39</sup>. Sam responds to this by offering as candidate solution an explanation of what to do in the task (ll. 6-7). Having co-established the purpose of the task with Sam (ll. 8-9), Valerio uses the MWE three times for a purpose which has not previously been explored in this paper; he uses it to ask for Sam's opinion on the first example. In l. 11, Sam repairs Valerio's turn and Valerio repeats the MWE as a candidate solution (l. 12). Then, in l. 13, Sam treats the MWE as a repair, seemingly offering an explanation of the task as candidate solution. In l. 14, underlining the co-constructed nature of the interaction, Valerio almost finishes Sam's turn, the conjunction *and* binding the turns in l. 13-14 as one unbroken syntagmatic strung (Lerner 1991) Eventually, l. 15, Sam reacts to Valerio's turn as Valerio intended; the intended meaning does not materialize until Sam reacts in an appropriate manner, and they get the task underway.

This interaction marks the second time, chronologically, that Valerio uses the expression in the sense of asking for another person's opinion. The first instance of this, as captured on tape, took place in the following interaction, approx. two months prior to Extract 8:

Extract 9 (mar 09 04)

1 Sara Valerio

2 Valerio ((gets up, walks to teacher)) () yes teacher

3 Sara okay () would you like to stay in level c or try level d

4 Valerio → ehm () I like level d but I don't know () what do you say ()

5 ((gesturing)) [mmm

6 Sara [how many times have you been in level c

7 Valerio two

---

<sup>39</sup> These two past tense uses are not counted as MWE in Table 1 but count as one separate type. This is merely a taxonomy issue and of no pivotal importance. Going by a purely interactional analysis, one might count them as MWEs.

- 8 Sara           just two times () three times is ok
- 9 Valerio       ((nods))
- 10               (...)
- 11 Sara         okay ((writes)) () how about you stay in level c one more time
- 12 Valerio     ((nods))

In this interaction, the focal student is summoned by the teacher to have a little talk on how things are going in class and which level he should most profitably attend next semester. It is worth stating here that in this ESL class, students attend levels A-D, A being beginning and D being intermediate (there are no advanced levels; see Brillanceau 2005; Reder 2005). At this stage the relevant discussion for our focal student is whether to advance from C to D. When asked by the teacher he says he wants to try level D. Then comes our focal expression *what do you say*, which is clearly not meant as a lexical inquiry of any sort in the way we saw in the preceding extracts. Also, it is not oriented to as such. The teacher immediately launches an inquiry pertaining to the topic of the sequence (namely Valerio's next level positioning) by asking him how many times he has done level C. He tells her 'twice' which she deems as insufficient for him; he is better off spending another term in level C which he ends up accepting (nodding in line 12 – note that the extract does not contain the whole sequence, l. 10). This usage of the MWE as a means to ask for other people's opinion in a non-repair way becomes successful, and perhaps entrenched, as shown in extracts 8 and 9 above and extracts 10 and 11 below. It is interesting, however, that Sam (extract 8) and the teacher (extract 9) react differently to this use of the MWE by Valerio. Perhaps because of the interactional history shared by Sam and Valerio in the classroom, Sam was prone to react to Valerio's MWE as a repair-initiator – until the time of Extract 8, this had, after all, been the typical job done by the MWE in interactions between Sam and Valerio. If this is a correct assertion, it is not unlike the finding in Brouwer and Wagner (2004: 35) that "participants' knowledge of each other builds the structure for future interactions".

Extract 10 May 14 04

- 1 Valerio       maybe ((pointing to two places in papers))
- 2 Dom          maybe okay that's a g that's a good thank you
- 3 Ian          oh maybe
- 4 Dom          may[be by:
- 5 Ian                 [okay in my opinion

6 Dom in my op yes:((reaches towards Ian))  
7 Valerio [yeah  
8 Ian [yes  
9 Dom yes  
10 Ian yeah in my opinion. [that  
11 Valerio [good! ((pointing at Ian))  
12 Ian that's bee ((pointing to Valerio's papers))  
13 Valerio what is your opinion?  
14 Ian heh in my opihinion ((laughing voice))  
15 Valerio what is your opinion. three hundred thou:sand?  
16 Ian it's u:h the elderly population is increasing b[y:  
17 Valerio [a month  
18 Ian by  
19 Valerio a month  
20 Ian three [hundred  
21 Valerio [maybe three hundred  
22 Ian [thousand  
23 Valeri → [what do you say? () three hundred thousand?  
24 Ian I guess  
25 Valerio okay

In this interaction, the students are discussing by how many individuals the world's elderly population is increasing. This extract starts in the middle of a long interaction in which Ian has shown his reluctance to answer the question, saying to both the teacher and Valerio that he does not know and that he is incapable of guessing. We enter the interaction at the point where Valerio proposes *maybe* as a means to initiate giving an estimate; at least that is how the co-participants respond to Valerio's *maybe* in l. 1. The teacher supports Valerio (l. 2) and Ian, in turn, changes his orientation to the task at hand in l. 3, *oh maybe*; the change is evidenced by the *oh* (Heritage 1984; quoted in Hutchby and Wooffitt 1998: 129). Ian then complies by offering *in my opinion* as a possible means to express his estimate, which receives positive assessments from both the teacher and Valerio (ll. 6-11). In l. 12 Ian actually gives his opinion on the matter, namely that 'answer B' in the task is the right one. For some reason, Valerio does not orient to this and goes on to specifically

asks for Ian's opinion (l. 13, l. 15) and later, using the MWE (l. 23), repeats the request for Ian's opinion on the matter. Interestingly, Valerio performs the first request for Ian's opinion in l. 13 by using the phrase *what is your opinion?*, which has been afforded (van Lier 2002) by the interaction. He repeats the phrase in l. 15 in a more quiet voice and with slightly falling intonation, implying that it might be private speech (Ohta 2001). They eventually end up agreeing, in a highly co-constructed manner (ll. 15-24) on the number 300,000 (which is one of three options given in the task). Towards the end of this co-constructed interaction, Valerio for the last time here asks for Ian's opinion, this time using the MWE; the use of the MWE even here, in an interaction which previously afforded the useful native-like phrase *what is your opinion*, suggests its entrenchment as a means to ask a for co-participant's opinion<sup>40</sup>.

In Extract 11 below, the MWE is yet again used to ask for somebody's opinion. The students are doing a task in which they discuss the contents of a story they have read in class. Valerio asks a peer for his thoughts on what one of the protagonists in the story was doing in Asia, and the peer responds relevantly (ll. 1-2).

Extract 11: 16-July 04<sup>41</sup> (group of three; Valerio, Leonid, Female student (TBD))

- 1 Valer → what do you say () Leonid () what was Arthur Burden doing in Asia?  
 2 Leonid xxx Arthur Burden take photograph  
 3 Valer he take the photograph? () he was[: photograph xxx microfilm ]  
 4 Leonid [phot\_ ((points to paper)) xxx] film  
 5 TBD I think maybe take this photographer xxx something  
 6 (3)  
 7 Valerio what did Arthur Burden take photograph of () the city?  
 8 ((someone behind them starts talking to Valerio and eventually the teacher terminates  
 9 group work))

Concluding the data-analysis section, and parallel with Extracts 6 and 7b+c, Extract 12 displays the use of the MWE as an other-initiated self-repair. As such, its use is trouble-free as the co-participant

<sup>40</sup> Another interesting observation in relation to Valerio's *what do you say* is that Ian actually, a few turns later, repeats it to ask for Valerio's opinion.

<sup>41</sup> There was one more instantiation of the MWE in this sense a few turns earlier in the interaction, but at that time, Leonid's response was inaudible.



treats it that way, too. This extract has been included to complete the data-set; all uses of the MWE have now been documented in interactional terms (except otherwise noted).

Extract 12, 30-June 05

- 1 Valerio so you work in different places you have idea you know other jobs  
2 May hh[: ((in-breath))  
3 Valerio [only one exactly. because I'm work only for housekeeping I don't work for other  
4 other job. I don't know I [don't have idea ((shakes head))  
5 May [but that's in my country though it's not here only xxx I only  
6 work at this tea shop  
7 Valerio tea shop.  
8 May yeah  
9 Valerio → tea shop is the is the is eh the what do you say before?  
10 May it's like a coffee shop?  
11 Valerio it's a restaurant? [coffee shop?  
12 May [yeah yeah yeah yeah it's kind of coffee shop you make drink?  
13 Valerio the same [for the starbucks and you can: [what do you drink  
14 May [xxx [yeah xxx like that yeah yeah yeah

Summing up, the MWE was initially used in lexical enquiry sequences. The orientation to the MWE in these cases by the co-participant was found to be dependent on whether or not it was accompanied by a summons or an active physical presence, primarily via eye gaze. If not, it was treated as a display of 'doing thinking'. These uses could be found in the data from July 03 through March 04 (extract 1-5). In November 03, Extract 6, a new use of the MWE emerges, namely as an other-initiated self-repair, as Valerio starts using it to ask for, or even challenge, something previously done or uttered by the co-participant – or in the case of 7a, perhaps something *not* done by the co-participant. This form of usage is found in the data from November 03 through July 05 (Extracts 6, 7a-c, and 12). Spring 04 marks the emergence of the other form of other-oriented usage for which the MWE is successfully used, as Valerio uses it to ask for other people's opinion on whatever task is at hand. This usage is found in the data from March through July 04 (extracts 8-11).

### Discussion:

Displaying how Valerio accumulates linguistic resources, the interactions represented here have pointed towards a principle of dual routinisation of action sequence and linguistic MWE, comparable to previously mentioned results of studies in language socialization (e.g., Kanagy 1999; Zuengler and Cole 2005). Arguably, what has happened, is that Valerio has been socialized into the activity of inviting for help (Brouwer 2003) in the classroom, specifically doing a lexical inquiry, using the particular MWE under investigation here. This use was demonstrated in extracts 1, 2, 3b, and 3c. As such, this also study confirms the findings in a current study (Author and Colleague2 2007) on business phone conversations between a spare parts representative and his business partners. The data for that study, however, covering a period of 4 years, showed a link between linguistic and interactional routines that did not develop at all – there was perfect standstill in this respect over 4 years – whereas these data imply that the MWE, at some point in development, spills over, as it were, into environments that are increasingly different from the one in which it emerged.

The first different environment is similar to the initial one in so far as it carries traces of doing a lexical enquiry; however, the interactional load it carries changes substantially as Valerio uses it to display that he is doing thinking. This usage was demonstrated in extracts 2, 3a and 5. The principles of routinisation, therefore, seems to hang together a view of learning corresponding to the participation metaphor and the ideas and principles prevalent in language socialization studies, whereas the use of the expression in the new environment is dependent on a different kind of learning, namely one that sustains a view of interactional abilities as partially portable; that some aspects of linguistic behaviour are retained by the individual and carried across contextual barriers (Larsen-Freeman 2004), i.e., acquired. The psycholinguistic whole unit status of the expression, and thus the likelihood of its being portable, is supported by two usages, not included in the transcribed extracts<sup>42</sup>, "I understand what do you say" and "I don't understand what do you say". This whole-unit status might be one of the reasons why Valerio uses it often and is capable of transporting it into new interactional environments.

Furthermore, the present research, fine-tuned to the investigation of emergent performance-driven linguistic inventories, has pointed towards an embrace of interaction and learning as points on a

---

<sup>42</sup> Together with Leonid's inaudible turn already mentioned, these are the only two usages found in Valerio's data that are not transcribed here.

continuum rather than each other's opposition. In categorical terms it is impossible to distinguish at what point interaction comes to a halt and learning takes over – or vice versa, for that matter. It has been argued that the interaction in which the MWE is deployed initially is in large part a reflection of the nature of the language classroom; as such the MWE becomes routinised in a practice in the classroom. However, at some point – it is impossible to say when or where – the MWE becomes more generally deployable, as it were. The locally contingent aspects of classroom interaction seem to have left a trace (Atkinson 2007) in terms of changes in the linguistic inventory of my focal student. We might speculate that this is a result of high frequency of the MWE itself and changing needs for Valerio to indulge in lexical inquiry practices, but the point of the trace leaving, as it were, remains elusive.

The findings here also seem to be in perfect alignment with some of the implications in Hellermann (2007), in which a classroom learner of English L2 was found to pick up from a peer and recycle a specific utterance *I talk to you* in identical sequence positioning in subsequent interaction. Hellermann's study constitutes a window onto the dual routinisation of social activity and linguistic utterance (Kanagy 1999; Author and Colleague2 2007), supporting a view of learning as situated in participation. In Hellermann's study *I talk to you* recurred in task openings, an activity inherent to the language classroom as a community of practice. As such, it supports the findings in Hellermann (2006), reported on above. The MWE investigated here could also be seen as initiating an activity of the language classroom as community of practice, namely an invitation for help (Brouwer 2003). The parallels between Hellermann's results and those of the present study do not end here, however. Hellermann (2007: 92) goes on to speculate that "...strategies used and learned in one situation may be applicable to other situations", and this is precisely what the extracts presented here have shown. For the MWE *what do you say*, then, it seems that what was at one point routinised in one sequential environment of inviting for help, a practice in the language classroom, later came to be employed in quite different situations – but situations carrying traces of previous usage. In the two initially occurring environments it was used as a self-initiated self-repair, but in later environments it was used, first as an other-initiated repair then as a completely other-oriented activity of asking for somebody else's opinion.

The data, then, indicate that the MWE in time becomes available to Valerio in more than one practice, ultimately pointing to the need for a timeline distinction between performance and

productivity. The former is always local and everywhere social; the latter need not be. This ultimately calls for the need to invoke both the participation and the acquisition metaphors for L2 learning. Therefore, I aim to rethink the performance-competence dualism in terms of a time-scale where *performance* is always and everywhere socially situated in the here-and-now world of the speakers, and *competence* is emergent as ever-changing ontogenetic residues or sediments of linguistic experience. This implies abandoning the traditional research paradigm in SLA in which the scope of inquiry is the development of a morpho-syntactic linguistic system of competence to yield all new research questions, framed around the study of 'environmentally adaptive behaviour that leaves a trace' (Atkinson 2007), where the stuff of learning is thought to be that which is recycled over time and/or carried across contextual boundaries (Larsen-Freeman 2004, Author subm.). This allows for an investigative framework for L2 learning which acknowledges both social, co-constructed and individual experience to investigate how “grammar and social interaction organise one another” (Schegloff et al. 1996: 33).. However, it is inferred that L2 learning, and thus researching L2 learning, should be centered around issues pertaining to the improvement of interactional abilities ('performance') as well as the broadening and deepening of the quality of the residual linguistic experiences ('competence') in ontogenesis. This in turn means that both the learning as participation and the learning as acquisition metaphors (Sfard 1998) hold equal relevance – both as far as objects of research are concerned and as desirable objects of achievement for the individual L2 learner. In this respect we have an ideal situation in which the etic, researcher-relevant, categories conflate with the emic, i.e. participant-relevant, ones (Firth and Wagner 1997; Markee and Kasper 2004; Kasper 2007). As will be inferred, then, the most important dualism to overcome is the performance-competence distinction by way of viewing learning as both participation and acquisition, thus leading the way to a framework which embraces both etic and emic categories of analysis.

Language learning is thus centrally a matter of *interactional competence* as it has been developed in some L2 research as language L2 users/learners' ability to employ (co-constructed) abilities, e.g., linguistic expressions and sequential routines, in interactive practices (see Hellermann 2007 and references cited there). In my opinion, interactional competence is in nature closer to performance than competence (traditionally conceived) and since I aim to rethink these notions I believe I will also need to build on the notion of interactional competence rather than straightforwardly adopt it. I envision language learning as interactional competence development as doing things in a real world

and getting better at things one can already do ('**performance enhancement**'), while also gaining a richer experience, a more evolved repertoire or inventory of resources which transcend the moment ('**productivity enhancement**'). This latter operationalization seems to be in alignment with Lantolf and Thorne (2006: 17) who propose that "learning an additional language is about enhancing one's repertoire of fragments and patterns that enables participation in a wider array of communicative activities" and is perhaps best thought of in terms of learning as acquisition, whereas the former might most profitably be dealt with in terms of learning as participation. The line between the two, however, remains purposely indistinct, and so expanding yet again on Atkinson (2007), we must look at *experience which leaves traces* – as well as the traces themselves. It is, after all, hard to argue that performance enhancement does not somehow come to be cognitively anchored in the individual; in fact, this is exactly where the notion of routinisation of activities becomes relevant. And implicit in that notion is also a notion of movement through time (routinisation presupposes the passing of time) which resonates with the overall idea of reconceptualising the competence-performance distinction along a time-dimension. Because language use by necessity is situated, performance enhancement is thought of as always social, bound by context, and at the mercy of locally emergent interactional contingencies; the emergent inventory of resources, while probably constituted by the sediments of performance enhancement activities, need not be. In the case of the present data, all uses of the MWE are a matter of performance enhancement; however, as the use of the MWE becomes gradually expandable it spills over into productivity enhancement. This introduces an elaboration of Larsen-Freeman's (2004) memorable conceptualization of learning as that which is carried across contextual boundaries. *What do you say* is carried across contextual boundaries, time-wise, but only so to a certain extent content-wise; i.e., the expression is locally, interactionally contingent, it does not become relevant in any old conversational situation. The utterance schema that emerges from the MWE, namely the more general *do-schema*, while thus a sediment of those interactional contingencies, is much more generally applicable and not, in terms of use, dependent on a narrowly defined conversational setting. The utterance schema, therefore, is carried across both content-defined and time-defined contextual boundaries. This, in essence, marks the difference, fluid as it is, between 'performance enhancement' and 'productivity enhancement' as I have operationalised them here.

### Conclusions and perspectives:

Recent CA-oriented classroom research has argued that language practices of individual classroom learners change with increased participation in recurring social actions which serve as “site[s] for language development” (Hellermann and Cole, *forthc.*). What we see here expands on that. We have an action sequence in which some social action is carried out (doing a lexical inquiry), coupled with a fixed expression. This fixed expression is then over time extended to be used in other contexts. Now, linguistic behaviour may be social in nature, but sociality should not necessarily be given prominence over individuality. Rather, the two should be seen as mutually constitutive. These data do not support the idea that individual linguistic development is driven solely by social actions which afford new utterances and constructions, but show us that 1) certain expressions at certain points in time sit in certain environments, with sociality and interactional requirements informing the nature of the language used, and that 2) reused linguistic material (recurrent MWEs) may act as guides in introducing the participant to new social actions in which sense individuality and sociality are reflexive of each other. Thus, microanalysis can be said to throw light on the early emergence, as it were, of linguistic material ontogenetically, but seems unable to handle ontogenetical development in a larger time perspective. Microanalysis has a lot to offer in terms of learning as participation but when it comes to investigating the portable nature of linguistic items, these being primarily MWEs and utterance schemas (Author *subm.*), other analytical methodologies and frameworks are probably necessary. On a speculative note one might pose the question if early learning is more characterized by a correlation between environment and expression than advanced learning, but that is one for the future to take up. A closer scrutiny than undertaken here of the question, also raised by Firth and Wagner (2007), on the nature of the changing environments in which people expand on their linguistic resources develop their linguistic inventories remains at the core of any approach to longitudinal L2 learning studies which aim to take interaction seriously.

From the perspective of investigating language learning as a locally applied usage pattern and the application of that same pattern over time, the data have shown how the use of the target utterance was initially situated in a recurring environment but later expanded to be used in other environments as well. It was argued that the identical deployment of the utterance in comparable sequences over time, an initial routinisation of the MWE, was an example of Valerio having been socialized into the language classroom practice of inviting for help (Brouwer 2003). This, in turn, was seen to hang together with the learning as participation metaphor, whereas Valerio’s ability to

use the MWE in a new context was argued to be more profitably thought of in terms of the learning as acquisition metaphor. Based on his activities in the social world of the classroom practices, then, he improved his productivity. As such, it has been demonstrated how a full account of L2 learning needs to take both participation in social interaction and psycholinguistic notions of cognitive portability, linguistic resources, in terms of acquisition in to consideration. Furthermore, it was demonstrated how these learning issues fit nicely into a rethinking of the performance-competence distinction along a time dimension of local performance enhancement (routinisation) and general productivity enhancement which transcends the moment. Whether one considers either side of the long-standing social – cognitive SLA dispute to be more profitable for investigating L2 learning, this study has shown that neither side alone is able to approach the immense task of attempting an exhaustive account of the processes involved. The present study does not presume to have shown such an account either; it does, however, see metaphor dictatorship as an un-navigable road for future research in L2 learning. The assumptions and principles empirically substantiated here, it is my hope, might constitute a stepping-stone for further investigating the prospects of a new SLA eclecticism.

#### References:

Atkinson, D. 2002. Toward a sociocognitive approach to second language acquisition. In *The Modern Language Journal* 86/4: 525-545

Atkinson, D. 2007. Toward a sociocognitive approach to second language acquisition. Colloquium talk presented at AAAL, Costa Mesa, California, 21-24 April 2007.

Author. Submitted. Constructing another language – Usage-based linguistics in second language acquisition.

Author and Colleague. 2007. Are recurring multi-word expressions really syntactic freezes? Second language acquisition from the perspective of usage-based linguistics'. In Nenonen, M and S. Niemi (eds.): *Collocations and Idioms 1: Papers from the First Nordic Conference on Syntactic Freezes, Joensuu, May 19-20, 2007*. Studies in Languages, University of Joensuu, Vol. 41. Joensuu: Joensuu University Press

Author and Colleague2. 2007. What counts as evidence in longitudinal language learning.? Paper presented at AAAL, Costa Mesa, California, 21-24 April 2007.

Bardovi-Harlig, K. 2002 'A new starting point? Investigating formulaic use and input in future expression.' *Studies in Second Language Acquisition* 24/2: 189-198.

Block, D. 2003. *The social turn in second language acquisition*. Edinburgh: Edinburgh University Press

Brillianceau, D. 2005. Spontaneous conversation. A window into language learners' autonomy. In *Focus on Basics 8a*

Brouwer, C.E., 2003. Word searches in NNS-NS interaction: Opportunities for language learning? *Modern Language Journal*, 87, 4. 534-545.

- Brouwer, C. E. 2004. Doing pronunciation: A sequence of novice-expert interaction. In: Gardner, R. & J. Wagner *Second language talk*. London & New York: Continuum.
- Brouwer, C. E. and J. Wagner, 2004: Developmental issues in second language conversation. In: *Journal of Applied Linguistics*, vol.1, 1, 29-47
- De Bot, K., W. Lowie, and M. Verspoor. 2007 'A Dynamic Systems Theory approach to second language acquisition.' *Bilingualism: Language and Cognition* 10/1: 7-21.
- Ellis, N. C. 2002. 'Frequency effects in language processing – a review with implications for theories of implicit and explicit language acquisition.' *Studies in Second Language Acquisition* 24/2: 143-188.
- Ellis, N. C. and D. Larsen-Freeman. 2006. Language emergence: Implications for applied linguistics – Introduction to the Special Issue. *Applied Linguistics* 27/4: 558-589.
- Firth, A. & J. Wagner (1997): On discourse, communication, and (some) fundamental concepts in SLA research. In: *The Modern Language Journal* 81, iii, 285-300
- Firth, A. and J. Wagner. 1998. SLA Property: No Trespassing' *The Modern Language Journal* 82/1: 91-94.
- Firth, A. and J. Wagner. 2007. S/FL learning as a social accomplishment: Elaborations on a 'reconceptualised' SLA. *The Modern Language Journal* 91, 800-819.
- Goffman, E. (1981). *Forms of Talk*. Blackwell, Oxford.
- Goodwin, M. H. and C. Goodwin 1986. Gesture and co-participation in the activity of searching for a word. *Semiotica* 62: 51-75
- Hellermann, J. 2006. Classroom interactive practices for developing L2 literacy: A microethnographic study of two beginning adult learners of English. In *Applied Linguistics* 27/3: 377-404
- Hellermann, J. 2007. The development of practices for action in classroom dyadic interaction: Focus on task openings. in *the Modern Language Journal* 91/1: 83-96
- Hellermann, J. and E. Cole (forthc.): Practices for social interaction in the language learning classroom: Disengagements from dyadic task interaction.
- Heritage, J. 1984. A change-of-state token and aspects of its sequential placement. In Atkinson and Heritage (eds.). *structures of social action: studies in conversation analysis* Cambridge: Cambridge University Press (299-345)
- Hopper, P. J. 1988. Emergent grammar and the a priori grammar postulate in Tannen, D. (ed.): *Linguistics in Context*. Norwood: Ablex.
- Hopper, P. J. 1998. Emergent grammar. In Tomasello, M. (ed.): *The New Psychology of Language, volume 1*. Mahwah, NJ.: Lawrence Erlbaum.
- Houtkoop-Stenstra, H. 1994. De interactionele functie van zacht spreken in interviews [the interactional function of low volume in interviews]. *Gramma/TTT*, 3, 183-202
- Hutchby, I. and Wooffitt, R. 1998 *Conversation analysis*. Cambridge: Polity
- Kanagy, R. 1999. Interactional routines as a mechanism for L2 acquisition and socialization in an immersion context. *Journal of Pragmatics* 31: 1467-1492
- Kasper, G. 1997. 'A' Stands for acquisition. A response to Firth and Wagner. In *The Modern Language Journal* 81/3: 307-312.



- Kasper, G. 2007. Conversation-analytic perspectives on language development. Discussant paper presented at AAAL, Costa Mesa, CA. April 21<sup>st</sup>-24<sup>th</sup>, 2007.
- Koshik, I. 2003. Wh-questions used as challenges. in *Discourse Studies 5/1*: 51-77
- Langacker, R. W. 2000. A dynamic usage-based Model. In Barlow, M. and S. Kemmer (eds.). 2000. *Usage-Based Models of Language*. Stanford: CSLI.
- Lantolf, J.P. (2000): Introducing sociocultural theory. In: J.P. Lantolf (ed.): *Sociocultural Theory and Second Language Learning*. Oxford: Oxford University Press)
- Lantolf, J. P. and S. L. Thorne. 2006. *Sociocultural Theory and the Genesis of Second Language Development*. Oxford: Oxford University Press.
- Larsen-Freeman, D. 1997: Chaos/Complexity science and second language acquisition. In: *Applied Linguistics 18/2*: 141-165
- Larsen-Freeman, D. 2006. The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. In *Applied Linguistics 27/4*: 590-619.
- Larsen-Freeman, D. 2007. Reflecting on the cognitive–social debate in second language acquisition. In *The Modern Language Journal 91*: 773-787
- Lieven, E., H. Behrens, J. Speares, and M. Tomasello. 2003. Early syntactic creativity: a usage-based approach. In *Journal of Child Language 30*: 333-370.
- Long, M.H. (1997): Construct validity in SLA research: A response to Firth and Wagner. In: *The Modern Language Journal 81*, iii: 318-323
- Markee, N. and G. Kasper (2004), Classroom talks: An introduction *The Modern Language Journal*, 88, 4: 491-500
- Mondada, L. & S. Pekarek Doehler (2004): Second language acquisition as situated practice: Task accomplishment in the French second language classroom. In: *The Modern Language Journal*, 88, 4, 501-518
- Myles, F., J. Hooper, and R Mitchell. 1998. Rote or rule? Exploring the role of formulaic language in classroom foreign language learning. In *Language Learning 48/3*: 323-363
- Ohta, A. S. 2001. *Second language acquisition processes in the classroom*. Mahwah, NJ.: Lawrence Erlbaum
- Reder, S. 2005. The "Lab School". *Focus on Basics 8a*.
- Reder, S., K. A. Harris, and K. Setzler. 2003. A Multimedia Adult Learner Corpus. In *TESOL Quarterly 37/3*: 546-557.
- Sacks, H., E.A. Schegloff, & G. Jefferson 1974: A Simplest systematics for the organization of turn-taking for conversation. In: *Language*, vol. 50, no. 4: 696-735
- Schegloff, E. A. 1979. The relevance of repair to syntax-for-conversation. In T. Givón (ed.): *Syntax and semantics: discourse and syntax* (vol. 12: 261-299) New York: Academic Press
- Schegloff, E. A. 1992. Repair after next turn: The last structurally provided defense of intersubjectivity. In *The American Journal of Sociology*, 97/ 5: 1295-1345
- Schegloff, E. A. 1996. Turn organization: one intersection of grammar and interaction. in: Ochs, E., E. A. Schegloff, and S. A. Thompson (eds.): *Interaction and Grammar* Cambridge, Cambridge University Press 52-133
- Schegloff, E. A. 2007. *Sequence organization in Interaction*. Cambridge: Cambridge University Press

- Schegloff, E. A., E. Ochs, and S. A. Thompson. 1996. Introduction. in: Ochs, E., E. A. Schegloff, and S. A. Thompson (eds.): *Interaction and Grammar* Cambridge, Cambridge University Press 1-51
- Schmitt, N. and R. Carter. 2004. Formulaic sequences in action: an introduction In Schmitt, N. (ed.): *Formulaic Sequences*. Amsterdam: John Benjamins.
- Sfard, A. 1998. On two metaphors for learning and the dangers of choosing just one. In *Educational Researcher* 27/2: 4-13
- Ten Have, P. 1999. *Doing conversation analysis*. London: Sage
- Thelen, E. and E. Bates. 2003. Connectionism and dynamic systems: are they really different? In *Developmental Science* 6/4: 378–391.
- Tomasello, M. 2000. First steps toward a usage-based theory of language Acquisition. In *Cognitive Linguistics* 11/1-2: 61-82.
- Tomasello, M. 2003. *Constructing a Language*. Cambridge: Cambridge University Press.
- Van Lier, L. 2002: An ecological-semiotic perspective on language and linguistics. In: Kramsch, C. (ed.): *Language Acquisition and Language Socialization*. (London: Continuum)
- Wagner, J. and Gardner, R. 2004. Introduction. In: R. Gardner & J. Wagner (eds.): *Second Language Conversations*. (London: Continuum)
- Watson-Gegeo, K. 2004. Mind Language and epistemology: towards a language socialization paradigm for SLA. In *The Modern language journal* 88/3: 331-350
- Wray, A. 2002. *Formulaic Language and the Lexicon*. New York, NY: Cambridge University Press.
- Young, R. F. and E. R. Miller. 2004. Learning as changing participation: discourse roles in ESL writing conferences. In: *The Modern Language Journal*, 88, 4: 519-535
- Zuengler, J and K. Cole. 2005. Language socialization and second language learning. In: Hinkel, E. (ed.): *Handbook of Research in Second Language Teaching and Learning*. Mahwah, NJ: Lawrence Erlbaum
- Zuengler, J. and E. R. Miller (2006). Cognitive and sociocultural perspectives: two parallel SLA worlds? *TESOL Quarterly* 40-1: 35-58

APPENDIX	Say	Repeat	Like	Ask	Write	make	Work	Use	Have	Do	Read <sup>43</sup>	Take	Own	Drink	Know	Pay	Think	live
<b>SUMMER 2003:</b>																		
<b>What do you v</b>	5																	
Do you v <sup>44</sup>		1	4															
How do you v	2																	
Why do you v			1															
<b>AUT 03</b>																		
<b>Do you Verb</b>	1		1	1														
<b>What do/did you v<sup>45</sup></b>	2				1 (did)													
<b>WINTER 04</b>																		
Do you v				1		1												
What do you V	2																	
Why do you v	1																	
What did he v	1																	
<b>SPRING 04<sup>46</sup></b>																		
what do / did you v	12 (2 did)				6 (1 did)			3										
Do you v							1											
When do you v								3										
<b>SUMMER 04<sup>47</sup></b>																		
<b>What do / did you</b>	2									2 (did)								
<b>Do you v</b>			2						6		1	1						
<b>Do 3<sup>rd</sup> pers pl</b>									1									
<b>SUMMER 05:</b>																		
Do you v	1		1				1		1						1			
What do you v	1		1					1	1	1				1			1	
What do 3 <sup>rd</sup> pers sg																2		
Where do you v																		3

Table 2: emergence of aux-do patterns. Please note that some patterns have been conflated (e.g. do and did verb) to save space.

<sup>43</sup> Valerio says *reading*.

<sup>44</sup> One afforded use not mentioned in the Table: 'do you dancing tango' Aug 12

<sup>45</sup> Nov 4: one recorded partially inaudible use not included in the table; it sounds like Valerio says: 'what do you quest'.

<sup>46</sup> May 4: One use not included: 'How long do you have exercised'. It a sentence Valerio has written in a task practicing present perfect; he asks his partner if it is correct.

<sup>47</sup> June 25: two uses that are counted but not included in the Table: *do* used to emphasise main verb: *When you do have...* and *I do remember...*

## Chapter 5: ESK4

### You no oich – user-based L2 learning: the case of negation<sup>48</sup>.

#### *Abstract:*

In this paper I discuss the application of Usage-Based Linguistics (UBL) to an analysis and description of developmental issues in second language (L2) learning. Building on previous and current research (Author and Colleague 2007; Author subm. a, subm. b, subm. c.), I draw on the UBL framework to argue that L2 learning is item-based and emergent, i.e., taking concrete linguistic material in the form of recurring multi-word expressions, as its starting point and slowly moving towards a dynamic, ever-changing inventory of linguistic resources as the language learner engages in different kinds of social interaction, 'usage events', over time. I will present results from on-going investigations into emergent linguistic inventories of Mexican-Spanish speaking learners of English in a classroom setting in the US. The focus is on the emergence of the *do-negation* construction and the results, indicating substantial individual differences, are not immediately compatible with proposed developmental sequences for negation (e.g., R. Ellis 1985/1996; VanPatten and Williams 2007). It is argued that because recurring interactional environments and linguistic development are inseparable, the major reasons for the individual differences are found in environmental aspects of the learning trajectories, pointing towards the fruitfulness of applying a locally contextualized apparatus to the description and analyses of L2 development (Author subm. a).

#### 1. Usage-Based Linguistics, Multi Word Expressions, and L2 learning.

Originally coined by Langacker as a super-ordinate cover-term for a number of functional-cognitive linguistic theories, UBL shares a number of key principles in approaching the study of language (e.g. Langacker 1987, 1988, 1991, 2000; Tomasello 1998; Barlow and Kemmer 2000; Tummers et al. 2005). Among them are the assumptions that language is part and parcel of general cognitive structures and capabilities, and that the structure of grammar has its source in language use. This

---

<sup>48</sup> Under revision, this paper is a reworked version of a paper delivered at the symposium "Cognition Applied: Usage-Based Linguistics and L2 Learning", University of Southern Denmark, August 16<sup>th</sup> 2007. I thank Teresa Cadierno for insightful comments and invaluable assistance in the work on the quantitative aspects of the data analysis. Needless to say, any flaws are my responsibility.

makes for a view of language knowledge that does not require a specialized organ, e.g. Universal Grammar, to ensure its successful acquisition. In fact, the notion of successful acquisition becomes a gradable one in UBL because language structure is seen to emerge from concrete usage events (Langacker 2000). Thus, shaking the foundations of the competence-performance distinction, UBL does not posit any ideas along the lines of an ideal speaker-hearer in an ideal language community. Rather, the language knowledge of any speaker-hearer is seen as the frequency-biased compilation of linguistic routines encountered in real-time experience (N. Ellis 2002). Language use and knowledge are therefore each other's prerequisites; there is a constant flow of information oscillating between them as people engage in interaction in the real world.

This UBL view of language knowledge as experience (Tomasello 2000; N. Ellis 2002) has implications for the notion of acquisition. As has been argued forcefully by Firth and Wagner (1998, 2007) and demonstrated in Author (subm. c), there can be no fixed demarcation line between language use and language acquisition. We simply cannot tell, by looking at L2 data, when one stops and the other takes over, as it were. Keeping apart interaction and acquisition in the influential interactionist framework (e.g. Mackey 1999; Gass 2003), the uptake metaphor is arguably too narrow to fully capture the processes at work in an undertaking as complex as the learning of a (second) language. Linguistic items are rarely just picked up in one interaction and then used again in a later one at the free will of a language learning individual navigating alone in the world – and this process is substantiated empirically even more rarely (Gass 2003; Kasper 2004). Instead, this individual is not navigating on his own but is fundamentally dependent on the navigations of other individuals in a social reality of the here-and-now (Atkinson 2002). In other words, language knowledge is fundamentally dynamic and distributed across interactional contexts, rendering language learning a complex and non-linear process. An insight gradually gaining foothold in psycholinguistic SLA (e.g. Larsen-Freeman 2002, 2006; N. Ellis and Larsen-Freeman 2006; de Bot et al. 2007), this has been around for some time in more socially oriented SLA approaches (e.g. Firth and Wagner 1998) where language learning is usually conceptualised as a locally contextualized activity (Mondada and Pekarek-Doehler 2004; Hellermann and Cole *forthc.*).

In more concrete terms, UBL assumes language knowledge to be a structured inventory of symbolic units (i.e., form-meaning pairings) along a continuum of linguistic expressions and constructions ranging from the totally specific to the maximally general (Langacker 1987, 2000; Achard 2007).

All linguistic units are defined and characterized fundamentally identically in this 'grammar' (Croft and Cruse 2004), the only difference being one of degree of underlying schematicity. In terms of language acquisition the continuum of schematicity is operationalised as a developmental trajectory going from formulas (*I dunno*) via partially schematic, partially concrete patterns (*I don't Verb*) to fully abstract constructions (*NP AUX NEG VERB*) (e.g. N. Ellis 2002; Tomasello 2003). In my current UBL for SLA framework (Author and Colleague 2007; Author subm. a, subm. b; subm. c), formulas are conceived of as recurring multi-word expressions (MWEs). A MWE is a recurring string of words used for a relatively coherent and stable communicative purpose. This operationalization is based strictly on recurring units in the data – i.e., no MWEs are defined a priori – and all references to issues of storage are avoided because of the assumed non-division between syntactically generated language and whole-unit storage. The patterns which are partially concrete and stable, and partially schematic and open, are operationalized as 'utterance schemas, borrowed from Tomasello (2000) to emphasize that some schematic knowledge is seen to sanction the use of given lexical items in the open slot. The posited level of ultimate abstractness consists of schematic knowledge of symbolic units; e.g., *I don't Verb* is specified more generally in the inventory as the negation construction *NP aux neg Verb*. This level of schematicity is reached in usage as the learner encounters and learns a sufficient number of instantiations of the pattern to be able to analogically form abstractions based on what is generic to the pattern. The maximalistic nature of the linguistic inventory, however, guarantees that what is learned as specific (e.g., MWEs) is not replaced over time by the more abstract constructional knowledge acquired; rather, abstract patterns and their specific instantiations may cohabit in the grammar (Langacker 2000; Achard 2007).

In terms of ontogenesis, the path of acquisition as presented in UBL is an empirically valid alternative to the traditional Interlanguage idea that there is a more or less predestined order of acquisition; the item-based path of language learning predicts incremental developmental movement in which concrete patterns link to previously experienced concrete patterns, in time, with enough exemplars of a given construction in place, sanctioning the emergence of abstract language knowledge in the form of schematicity. L2 learning, in this vein, is not researchable along the lines of cross-constructurally deployed morpho-syntactic rules, as it is highly unlikely that this is in fact how people learn a L2 (Author subm. b). Rather, what people seem to be learning is a patchwork of interrelated linguistic patterns that are more or less schematic in their underlying representation, as put forward by the UBL inventory continuum; for L2 learning the degree of schematicity is always

an empirical issue and will depend on the linguistic pattern under investigation. Arguably, the least schematic and hence most concrete, i.e., lexically filled, patterns is the best starting point for analysing and describing L2 development (Author *subm. a*). This is essentially different from the traditional conception that learners go through certain stages of syntactic development, often called 'developmental sequences', when acquiring a language, an idea which has been very influential in SLA (see e.g., Meisel *et al.* 1981; Pienemann *et al.* 1988; Larsen-Freeman and Long 1991; R. Ellis, 1994; Glahn *et al.*, 2001; Pienemann, 2003; Håkansson 2005; Unsworth *et al.*, 2006). Although some researchers have been sceptical regarding the universalism implied by the predestined nature of the developmental sequences (e.g. Lund 1996, 1997), the fundamental view of L2 development as the gradual improvement of mastery of morpho-syntactic structures of the ever systematic Interlanguage has remained relatively unchallenged in psycholinguistic SLA research<sup>49</sup>.

Instead, refusing to accept the idea of a visible point of completion of the target language structure along the traditional Interlanguage route of acquisition (Firth and Wagner 1997), I propose an emergentist outlook on development. For the purpose of this study, emergentism implies two things: one is programmatic (and also substantiated by data): I pose no end-point of acquisition, the assumption being that spoken interaction is a source of constant renewal of the 'grammar' (i.e. language knowledge), implying infinity of acquisition in ontogenesis. The other is reflected in the data (which I will return to) where the emergent nature of the expressions, patterns and constructions is seen in the way that these act as each other's stepping stone in ontogenesis, as it were. This also underlines the programmatic statement, reiterated above, that MWEs are not beyond the current linguistic capabilities of the learner; rather, they should be seen as integral to the emergent linguistic resources as such.

## 2. Literature review: UBL in SLA.

Applying the UBL trajectory of L2 learning as it is proposed by Tomasello (2003) for L1 learning is a very recent undertaking. In fact, Author and Colleague (2007) were the first to chart Interlanguage

---

<sup>49</sup> In December 2006, however, in a special issue in Applied Linguistics focussing on emergentism and SLA, Larsen-Freeman criticized this traditional Interlanguage view of development, according to which target languages are viewed as stable and homogenous, Interlanguages are rule-governed, systematic and predictable and variance therefore uninteresting, and language in general is viewed as a purely cognitive. Instead, Larsen-Freeman, advocating a view of language as both social and cognitive, sees language learning as more complex than that; it is, she argues, to be understood as fundamentally dynamic, constant, and un-ending, with variance as something to take seriously (Larsen-Freeman 1997, 2002, 2006; N. Ellis & Larsen-Freeman 2006). Taking these fundamental stances to heart this paper questions the validity of applying an essentially timeless and rigorous theoretical framework to something that is fundamentally dynamic and distributed (Larsen-Freeman 2006; Author *forthc. a*).

development along those lines. Previously, the only attempt at applying these UBL developmental insights was Bardovi-Harlig (2002). In her examination of future tense expression by 16 learners of English, however, she operationalised formulas as is traditionally done syntactocentrically, by defining them as being beyond current Interlanguage capabilities. In a genuine UBL setting, her starting point would not have been an Interlanguage grammar that separates some utterances from the main grammar machinery in a periphery-core vein. In fact, the UBL view of development as explicitly exemplar-based precludes the separation of such ontogenetically initial exemplars (i.e., ‘formulas’) from being shunted to the muddy backwaters of a grammar periphery. Furthermore, UBL programmatically states that it is the language user rather than a language system that operates in language use (Langacker 2000). Mistaken starting point aside, Bardovi-Harlig (2002) found two phenomena indicating that the UBL path of acquisition is only partially valid for SLA: 1) the role of formulas in initial development was found to be limited; and 2) the use of formulas was found not to diminish ontogenetically. The first phenomenon, Bardovi-Harlig concedes, may be due to the fact that initial formulaic use is too brief to be detected in her corpus. In terms of the second one, Bardovi-Harlig fails to acknowledge that the existence of the formulas in advanced stages of learning in fact supports the validity of the suggested path of acquisition insofar as it is connected with the cohabitation in the grammar hypothesis. Bardovi-Harlig does conclude, however, that N. Ellis’ proposed path of acquisition presents a richer view of SLA development than a starting point that excludes formulas.

In a more recent study of the development of *do-negation* by a Mexican learner of English, parts of which are replicated here, Author and Colleague (2007) found positive evidence for the UBL path of acquisition, with the system emerging in acquisition characterized as the gradual abstraction of regularities that link expressions as constructions. *Do-negation* learning was found to be initially heavily reliant on one specific instantiation of the pattern, *I don’t know*, with productivity gradually increasing as the underlying knowledge seemed to become increasingly abstract, as reflected in type and token frequencies. *I don’t know* was also found to be stable throughout development suggesting its entrenchment as a MWE.<sup>50</sup> Their findings thus suggest that L2 learning is indeed item-based, that expression entrenchment is dependent on token frequency (as in the case of *I don’t know*) and that more abstract pattern and construction learning is dependent on type frequency, as expected.

---

<sup>50</sup> Entrenchment refers to the idea that linguistic items may be routinized through frequent use (Langacker 2000).



In a current longitudinal study on the same student, this time focusing on *can*-patterns, Author (subm. a) found initial learning to be dominated by a few concrete interrelated patterns, such as *I can write* and *can you write*, supporting the notion of L2 learning as item-based. However, no conclusive evidence was found that in L2 development language knowledge becomes increasingly abstract; patterns could not be shown to be linked as abstract constructions in ontogenesis. Instead, this study found that certain MWEs were inextricably linked with certain interactional contexts, suggesting a locally contextualized view of L2 learning as indistinguishable from L2 use, and showing interaction to be a constant source of renewal for the individual linguistic repertoire. Productivity in language learning, the study suggested, seems to be guaranteed by a rich inventory of linked utterance schemas rather than the learning of increasingly complex or increasingly schematised linguistic knowledge. Given this lack of schematicity the study suggested that L2 ontogenesis is best captured in the least abstract linguistic terms possible. In current research, Author (subm. b, subm. c) has found further evidence that the bulk of portable linguistic experience consists of utterance schemas which initially in development are dependent on a few, sometimes locally recurring, interrelated specific instantiations.

The present study explores the coming together of linguistic and interactional development. Focussing on the development of *do-negation*, I superimpose the method of type and token analyses from Author and Colleague (2007) on another focal student and discuss the picture of L2 development surfacing from this procedure against the backdrop of traditional Interlanguage studies. The ensuing differences between the two students call for a look in more detail at the actual usage events; i.e., the real-life occurrences of language in action where the linguistic inventory takes shape in L2 learning. I contextualize the trajectory of the new focal student's linguistic resources in ontogenesis as they co-emerge alongside aspects of interactional development, arguing that the emergence of certain patterns at certain points in time may be an artefact of classroom activities.

#### 4. Data discussion:

The data source for the present study is the Multimedia Adult English Learner Corpus (MAELC),<sup>51</sup> which consists of audio-visual recordings of classroom interaction in an English as a Second

---

<sup>51</sup> MAELC was compiled and is maintained at The National Labsite for Adult ESOL (known locally as the Lab School). The Lab School was supported, in part, by grant R309B6002 from the Institute for Education Science, U.S. Dept. of Education, to the National Center for the Study of Adult Learning and Literacy (NCSALL) and was a partnership

Language (ESL) context. The classrooms, in which the recordings were made, were equipped with video cameras and students were given wireless microphones on a rotational basis; the teacher also wore a microphone (Reder et al. 2003; Reder 2005). Consisting of recordings from July 2003 through July 2005 and September 2001 through February 2005, respectively, the present research is based on the study of two students, Valerio and Carlos<sup>52</sup>, adult Mexican male learners of English whose L1 is Spanish. The final database of the inquiry consists of transcripts from approx. 70 sessions (each consisting of three hours of recordings, not all of which has been transcribed) in which Valerio and Carlos are either wearing a microphone or sitting next to someone wearing a microphone.

#### 4.1 Carlos:

In Author and Colleague (2007), we observed three distinct but related kinds of negation pattern in the totality of Carlos's data:

- a) recurrent target language (TL) MWE: *I don't know*
- b) learner pattern: *Subj no V* (e.g., *I no remember*)
- c) TL pattern: *do-neg* pattern (e.g., *I don't think so*).

To get an overview of Carlos's inventory, we counted all instantiations (i.e., tokens, to be explained below) of *do*- and *no*-negated patterns in the data. The result of this count is displayed in Figure 1 below. Roughly, Carlos's development is characterised by three phases: in 2001, negation patterns consist of 1/3 MWEs and 2/3 learner patterns (2001), in 2005, we see a 42%-58% division between MWEs and other TL patterns, and in between there are varying stages of competition between the learner pattern and the TL pattern, with the MWE always present, fluctuating between 25% and 65% of total negation use. We thus noted three tendencies: long-term presence as well as initial acquisitional importance of the MWE, increase in TL pattern usage, and decrease (until the point of disappearance) in learner pattern usage<sup>53</sup>.

---

between Portland State University and Portland Community College. I am extremely grateful to Steve Reder and all the staff at The Lab School without whom this research would not be possible.

<sup>52</sup> Carlos was the focal student in the original study, Valerio in this one. Both are pseudonyms.

<sup>53</sup> Because of the on-going nature of the investigations reported on here and in the previous study we have found more instances of negation in Carlos's data. None of these new ones, however, undermine the findings of the previous study. I have appended an updated count of Carlos's negation. In the following brief outline of Carlos's data, the figure and tables are all taken from Author and Colleague 2007. I refer to the original article for in-depth analyses.

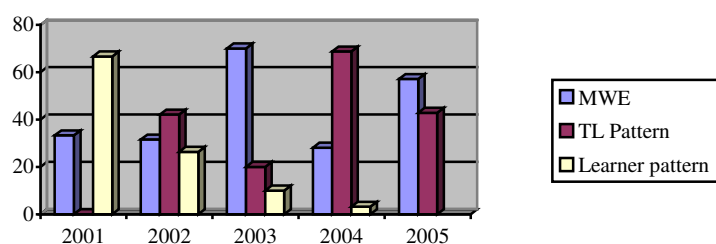


Figure 1: Carlos's negation pattern development

To investigate the possibility of an increased schematicity in Carlos's *negation-pattern* representation, we used type and token frequencies as analytical tools. As mentioned, Figure 1 is based on a token count. Tokens are the number of actually occurring concrete expressions, i.e., specific instantiations of the pattern under investigation, whereas types are the number of different instantiations of the same pattern. The interesting numbers are the ratios, however. A high token frequency and a low type frequency results in a low type-token ratio implies the existence of one or more MWEs, i.e., possible entrenchment of specific exemplars, whereas a high type to token ratio implies a high degree of productivity of the pattern, indexing possible schematicity of underlying representation. Even though development did not seem to be strictly linear, we did find a general increase in type-token ratio over time, suggesting that Carlos's TL pattern was becoming increasingly more abstract in its schematic representation as productivity is expanded in terms of negated verbs. The tendency is displayed in Table 2, the non-linearity evident in the equally high ratio numbers for 2002 and 2005.

	TL Pattern			Learner Pattern		
	Token	Type	Ratio	Token	Type	Ratio
2001	5	1	--	10	8	0,80
2002	14	7	0,50	5	4	0,80
2003	9	3	0,33	1	1	--
2004	31	13	0,42	1	1	--
2005	14	7	0,50	0	0	0

Table 2: Type and token frequencies for Carlos's TL pattern and learner pattern

This movement towards increasingly abstract knowledge was confirmed by an expansion in the combinatorial possibilities in the pattern per se, as reflected in increasingly varied use, not only in the Verb slot but also in the Subject slot, pronoun-filled. Table 3 below shows what patterns are in use in what period with what pronouns. While supporting the general developmental picture as

portrayed by the type-token count discussed above, Table 3 also points to the secret behind the high type-token ratio found in 2002. The internal variation for the TL pattern in 2002 was limited to two interrelated utterance schemas, *you don't verb* and *I don't verb* rather than a fully abstract construction, which was not found to emerge until later. In other words, the type-token ratios did not necessarily tell the full story of productivity; to get at this, it was pertinent to go beyond the numbers to look at the linguistic content in the types and tokens. Having done this, we found the linguistic variation needed to posit the existence of fully schematic linguistic knowledge at the end of the final two recording periods.

	2001	2002	2003	2004	2005
1 <sup>st</sup> pers., sing.	TLP / LP	TLP	TLP / LP	TLP / LP	TLP
2 <sup>nd</sup> pers., sing.	LP	TLP / LP			TLP
3 <sup>rd</sup> pers., sing.	LP	LP			TLP
1 <sup>st</sup> pers., pl.				TLP	
2 <sup>nd</sup> pers., pl.					
3 <sup>rd</sup> pers., pl.				TLP	
Past tense				TLP	TLP

Table 3: Carlos's negation pattern abstractness development. Note: TLP=Target Language Pattern; LP=Learner Pattern; pers.= person; sing.= singular; pl=plural.

A system was thus seen to be emerging and the learner pattern seemed to be declining. Conflating the combinations in use in the final two recording periods results in an almost full representation of combinatorial possibilities in terms of the *do-neg*-pattern which led us to posit the existence of an abstract underlying *subj do-neg verb*-construction.

#### 4.2. Valerio.

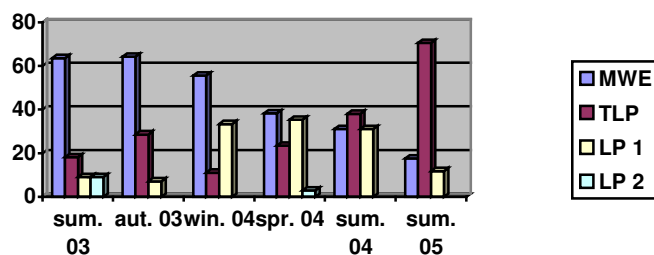
The data for Carlos were thus found to support the view of L2 development as item-based and so out of natural curiosity and encouraged by this success, I went on to conduct an identical analysis on negation pattern development in another student, Valerio. At first, I note that his linguistic inventory in totality for the investigated pattern looks both similar to and different from that of Carlos:

- a) Recurrent Target Language (TL) MWE: *I don't know*
- b) TL pattern: *Aux neg*-patterns (*I don't think*)
- c) Learner pattern 1: *Subj no V*-patterns (*I no remember*)
- d) Learner pattern 2: *Subj am no V*-patterns (*I am no say*)

As was the case with Carlos, the negation inventory consists of a number of interrelated patterns, these being the MWE *I don't know*, the target language pattern (*do neg*), and the learner pattern (*subj no verb*). The most striking similarity is the co-presence of the MWE, other instances of the *aux neg* pattern, and the *subj no v*-pattern in both students. The learner pattern 2, however, is a different matter. This seems to be idiosyncratic to Valerio and seemingly inexplicable; it surfaces at different points in time and it is used with some of the same main verbs as the other patterns. Because of its overall low frequency, I will not look any further into this pattern at this stage.

Figure 2 below displays an overview of the negation patterns found in the Valerio data and how their usage frequencies develop over time. Some developmental tendencies seem identical in the Carlos and Valerio: the MWE is frequent and retained throughout development (though decreasing in Valerio's case, whereas it was fluctuating in Carlos's case). TLP usage increases over time and there are various levels of competition in-between the two extreme points in time

Figure 2: Valerio's negation pattern development



However, there are also individual differences: in Valerio's case LP1 expands along with the TLP which makes for a less linear and more complex (or chaotic, if you like) picture of language development. For example, Winter 04 marks a LP1 explosion, so to speak. Looking in more detail at this period, I found a local recurrence of the pattern *you no write*, recurring in specific usage events in a task environment, to be discussed in section 5 below. This variation, then, may be best thought of as a reflection of classroom activities; the phrase is prompted by the nature of the tasks in the classroom. So the usage explosion seems to be about the verb negated and the context that prompted it, hence locally contextualized. In terms of the competition between the two patterns, and as might be inferred from figure 2, Spring and Summer 04 display equal usage of LP1 and TLP respectively. There does not seem to be any stage-defining dominant form on the basis of which one might tell developmental sequences apart. There seems to be no clear-cut development going from

learner patterns to target language patterns. In order to get at this language learning issue, the next question is “which pattern is dominant in terms of productivity/creativity?”

	TL Pattern			Learner Pattern		
	Token	Type	Ratio	Token	Type	Ratio
Sum 03	9	3	0,33	1	1	--
Aut 03	13	4	0,31	1	1	--
Win 04	18	4	0,22	13	5	0,38
Spr 04	21	8	0,38	12	10	0,83
Sum 04	29	12	0,41	13	11	0,85
Sum 05	15	8	0,53 <sup>54</sup>	2	1	0,50

Table 4: type and token analysis of target language and learner patterns, Valerio.

Looking at Table 4, one notes the same tendencies as in Carlos (learner pattern 2 has been excluded from this analysis due to low over-all frequency):

- the TL pattern, it is suggested, is becoming increasingly more abstract in its underlying schematicity as creativity is expanded
- the use of the learner pattern wanes to the brink of disappearance.

A few things are different, though. Note first the increase in total numbers of types and tokens for both patterns until Summer 05, at which point the LP disappears and the TLP becomes totally dominant. The ratios interestingly suggest increasingly high productivity for both patterns (LP until 05 when it disappears). This was not the case with Carlos whose developmental road was less bumpy<sup>55</sup>, with the LP gradually losing weight and the TLP correspondingly becoming more widely used and productive. Furthermore, although employed less frequently in total numbers, the LP is

<sup>54</sup> Valerio uses *I don't have idea* and *I don't have time* at this stage. The former recurs four times in his speech, the latter does not recur. They have been counted as two types even though they contain the same the main verb, because the former seems to be an entrenched form-meaning pattern used by Valerio to express that there is something he doesn't know about the present topic of conversation.

<sup>55</sup> Given the locally contextualized nature of Valerio's linguistic development, which is explored here, it might also be the case that Carlos's more linear development was due to less changing speech environments. This remains largely speculative, but a point in that direction is the fact that the high number of tokens, but relatively low type-token ratio, in 2004 was because Carlos in that period used four recurring verbs, namely *have*, *think*, *remember* and *know*, suggesting that usage was bound to specific environments.

more creatively productive here than the TLP (indicated by the higher type-token ratio). This is also reflected in that the TLP is characterised by an overwhelming use of the MWE – in 2004, 40 out of a total 68 tokens (59 %) were MWEs – and a more scarce use of other more rarely recurring patterns that are of no immediate interest here. In contrast, the LP in 2004 has four recurring expressions, *you no write* which recurs five times, and in addition two that recurs twice and one that recurs three times, accounting for only 12 out of a total 34 tokens (35 %).

	Sum 03	Aut 03	Win 04	Spr 04	Sum 04	Sum 05
1 <sup>st</sup> pers., sing.	TL / LP	TL	TL / LP	TL / LP	TL / LP	TL
2 <sup>nd</sup> pers., sing.		LP	TL / LP	TL / LP	TL / LP	TL
3 <sup>rd</sup> pers., sing.				TL / LP	LP	
1 <sup>st</sup> pers., pl.						
2 <sup>nd</sup> pers., pl.						
3 <sup>rd</sup> pers., pl.					LP / TL	TL
Past tense					TL	

Table 5: Valerio, development of schematicity?

Table 5 shows what negation patterns are employed at what stages and with what PRNs and tense morphology. Note again the exclusion from this table of LP2 because of low overall frequency. Table 5 supports what was suggested by the type-token relationships earlier, namely that the learner is working on increasingly abstract language system. This is evident in the expansion of the combinatorial possibilities and the possibility of expressing past tense, which does not seem to be feasible in the learner pattern for either student. However, it will also be noted that at no point in the data is there total freedom of combinatorial possibilities, as it were – the 1<sup>st</sup> and 2<sup>nd</sup> person plurals are never employed, and the final period related here shows fewer such possibilities than were in use for both the LP and the TLP in the preceding period. For Carlos only 2<sup>nd</sup> pers plural was never used, and therefore, in Author and Colleague (2007) we felt more comfortable arguing for schematised knowledge in Carlos's case.

The issue of schematicity is a tricky matter, though. Elsewhere (Author subm. a), studying Carlos's development of *can*-patterns, I could not find empirical evidence in favour of increased underlying schematicity as learning proceeded. In that paper, I argued that the finer details of L2 development

cannot be investigated by recourse to the most abstract categories usually preferred by the linguist. Studying ontogenesis, therefore, should always assume the most concrete starting point possible as far as linguistic pattern learning is concerned. Then, it is an empirical issue to establish or posit the existence of underlying schematicity, depending on the pattern under investigation. In a different study (Author subm. b), in which I investigated the interplay between linguistic routines and creativity, I concluded that the bulk of the linguistic inventories of my focal students (also Carlos and Valerio in that study) consists of utterance schemas – i.e., part creative, part stable patterns such as *I don't verb*, and *how many x* – suggesting that ultimately abstract schematicity may be an exception to the general rule of pattern learning. Therefore, my preferred point of departure is the emergentist idea of structural non-completion of acquisition – that (conversational) interaction impacts and changes 'grammar' to such an extent that it is futile to speak of a 'final state of L2 acquisition'. Language knowledge, it is believed, is fundamentally transitory (Firth and Wagner 1998) and intrinsically linked with real world activities. When data suggest that language knowledge evolves in ways that support the idea of an increasingly abstract underlying knowledge, then this should be taken into account. But investigative starting points are most profitably thought of in terms of specific patterns with the issue of schematicity remaining to be empirically established.

#### 4.4. Implications for L2 developmental sequences

The development of negation pattern as found in both focal students, and thus the view of L2 development propagated so far in this paper, is in contrast with the general line of argumentation in research into the so-called developmental sequences, where it is argued in a classical Interlanguage vein that all learners pass through certain structurally systematic developmental stages on their way to mastery of the L2 syntax. For negation, these 4 stages are defined by dominant structures that will be delineated below (taken from R. Ellis 1985/1996) and backed by examples from the present data. Succeeding the taxonomy outline, the relevance of the sequences themselves will be discussed on the basis of the present data:

- 1) 'External negation'. Here, the learner puts the negation particle outside the utterance nucleus. From the present data, Valerio's use of *no is x* (Author subm. b) will serve as an example.
- 2) 'Internal negation' stage where the negation particle is put inside the nucleus; the learner, though, is still incapable of negating systematically in a TL fashion. For



English as L2, the learner uses a pattern like the one observed in the present focal students: *subj no verb*. The negation particle may alternate with *not* or *don't*<sup>56</sup>. In the case of *don't* usage at this stage, the particle is thought of as unanalysed by the learner because it is beyond the current interlanguage level. It is, in other words, the particle placement that counts as primary criterion for defining the stage, not the actual items involved.

- 3) 'Modal verb negation'. At this stage, usually the learner has moved on to negate *can*, *will*, *shall*, etc. all in a structurally native-like fashion. Carlos's *you can't write 'which'*, is an example from the data.
- 4) TL rule acquired. The theory predicts the TL rule to be fully acquired for all negated contexts at this stage. Both Carlos and Valerio show numerous of instances of TL-like negation, e.g., the *do-neg.* discussed above, but at no point in either student is the 'rule' evidenced to be followed for all combinations.

While the structural characteristics of all 'stages' as represented by the outlined developmental sequences are all to be found in the present data, something is not quite right. Starting with stage 1, Valerio's external negation pattern, as exemplified above, is quite productive; his usage including some variety in the x-position, e.g. *no is good*, *no is true*, but also *no is melon*. Ontogenetically, Valerio uses this pattern throughout development; it does not disappear and no linear movement from *no is* → *it's not* can be documented. In terms of the structure of the linguistic inventory, this also has implications. The transcendent existence of this pattern alongside the documented development of the *do-neg*-pattern, as portrayed in the previous section, supports the idea of a construction-based approach to language knowledge and learning. The two, while perhaps interrelated to some extent which would be beyond the scope of the present paper to discuss, seem empirically established to be two separate patterns in learning. Carlos, on the other hand, has no external negation pattern in his linguistic arsenal, thus questioning the validity of positing this as a dominant form characterising an acquisitional stage in the first place.

---

<sup>56</sup> Recently, VanPatten and Williams (2007) outlined the same sequences, dividing this stage into two with exclusive use of *subj no verb* constituting a phase in its own right. The following phase is then seen to be characterised by alternating use among *no*, *not* and *don't*. R. Ellis' version is chosen here because, even though the sequences themselves are fundamentally questioned here, his layout seems to fit the present data a little bit better than that of VanPatten and Williams.

In terms of the internal negation pattern, stage 2, the data presented here and in our previous study (Author and Colleague 2007) support the general consensus that such a pattern does not at any point solely constitute a learner's negation usage abilities (R. Ellis 1985/1996); rather, it is there as a pattern 'competing' with the TL pattern. The difference between the path of learning as portrayed here and that generally conveyed by traditional developmental sequences studies is that the present research focuses on the genesis and ontogenesis of singular patterns that are constructionally related but not identical, whereas more traditional research focuses on the emergence of a syntactically rule-governed positioning of a broadly applicable negation particle. This is evident in the fact that this latter tradition, as mentioned above, conflates the learner pattern *s no v* and the TL pattern *s don't v*, because the syntactic placement of the particle, irrespective of different surface realizations, is identical in the two patterns. The present research, on the other hand, focuses on the empirical fact that the patterns, though interrelated, display acquisitional trajectories of their own; neither the learner pattern nor the TL pattern can be generalised to all negated contexts, so 'negation' does not seem to be learned as a syntactic phenomenon to be deployed across diverse linguistic patterns in a broad-sweeping manner.

The third stage, dominated by the modal negation, does not seem to exist as such from the point of view of the present data. For Carlos the reality is that roughly 18 months into his biography as L2 learner in this classroom, he uses *can't* for the first time (Author subm. a), *you can't write 'which'*, as he is discussing a meta-linguistic issue with his partner in a task-solving situation. Carlos never negates *can* in any other way, and because he does not use other modal verbs in negated contexts, it is inferred that the pattern is learned as such in an item-based manner rather than as part of a broadly applicable syntactic rule<sup>57</sup>. For Valerio, things are a bit different. Similarly to Carlos his use of *can't* emerges almost two years after he started attending ESL class. Contrary to Carlos, however, his use of *can't* is preceded by a momentary use of *no can verb*, making Valerio a better match with the proposed developmental sequences than Carlos.

As far as reaching stage 4 is concerned, Valerio's and Carlos's respective learning trajectories fit with this prediction to varying degrees. Whereas the learning of *can't* by Valerio was shown above to better fit the proposed developmental sequences than Carlos's learning of the same pattern, in the overall picture, Carlos seems to tally better with a priori developmental predictions. Carlos shows

---

<sup>57</sup> The item-based learning of the pattern is also supported by the fact that the verbs used in the negated *can*-contexts are already in use by Carlos in other related *can*-patterns (Author subm. a).

no traces of any learner patterns at the end of the data collecting, whereas Valerio, while using TL negation in many contexts, seems to retain the *no is x*-pattern; a stage 1 negation structure according to the developmental sequences. While it could easily be argued that full TL competence would in fact be the ideal end-point of learning as far as the various negation patterns are concerned, it is not immediately confirmed as a straightforwardly defined ‘stage’ by these data due to cross-constructural variation; TL patterns are in use by both students throughout development.

Even though traditional Interlanguage research in developmental sequences does allow for learner variation stage overlaps (R. Ellis 1985/1996), the findings presented here are ultimately incompatible with the underlying *descriptive* taxonomy of that framework because of the empirical invalidity of any dominant forms along a predestined road of learning. Discrete steps of learning cannot be discerned in a cross-constructurally structural manner. Pattern learning is best conceptualised as happening along, rather than across, constructional lines. Related to this lack of isolable points of learning, the present findings also go against the *acquisitional* insights as well as the linguistic theory underlying traditional research in developmental sequences, because ‘negation’ does not seem to be a structural form that is learned irrespective of negated items and learned to be applied cross-constructurally in a broad-sweep manner. The posited robustness of the developmental sequences is probably in part a question of research methodologies. Studies in favour of the sequences are overwhelmingly quantitative, aiming for very general tendencies at a large scale at the cost of acquisitional insights. Once we look into the details of ontogenesis, however, those tendencies yield to a much more complex portrayal of L2 learning (de Bot et al. 2007).

In addition to this, I argue that because of the underlying theory of linguistic knowledge research of that kind is fundamentally misguided. In this paper, therefore, the relevance and validity of the structural implicational scale underlying the sequences is questioned. Having empirically validated a view of L2 learning as fundamentally item-based, I view language knowledge as probabilistic and emergent (and locally situated, which will be the focus of the next section). Therefore, any a priori systemic, stable, and omnipresent conception of language will not do to describe and analyze L2 ontogenesis. L2 learning seems to be much too (lexically) specific to abide by a priori defined decontextualized rules of syntax, and the processes involved, socially and cognitively, do not seem to be concerned primarily with the gradual approximation to a homogenous target language, as

assumed in Interlanguage theory. Traditional Interlanguage studies are concerned with ‘generically applicable’ morpho-syntactic forms – i.e., past tense *-ed* which is generically applicable to verbs, or *neg.*, generically applicable to all *aux neg*-patterns. I argue that it is more fruitful to work on a pattern-by-pattern basis to investigate what items are used in accordance with target-language conventions and how these items are related to the other parts of the linguistic inventory under investigation. Doing this, I also argue for a shift away from “timeless mental competence and atemporal structural analysis” and instead embrace language use and acquisition as emergent, locally contingent, and “temporally and spatially situated” (Schegloff et al., 1996: 6). These aspects of local contextualization and situatedness of L2 learning is the focus of the next section.

### 5. Towards an embrace of a locally contextualized view of L2 development

This section is devoted to an investigation of the interactional side of linguistic development in a L2. It was noted above for Valerio, in the discussion on Table 4, that his type-token ratios suggested increasingly high productivity for both learner pattern (LP) and target language pattern (TLP); for the LP until 05 at which point its status has been degraded to the verge of disappearance. Furthermore, although employed less frequently in total numbers, Valerio’s LP, after having seen a veritable explosion in usage as mentioned, seemed to be more creatively productive than the TLP. This is an interesting learning issue; in traditional Interlanguage studies there is always a dominant form in all phases of development, but this does not straightforwardly fit Valerio<sup>58</sup>. Therefore, the focus of the present section is the use of the LP in the three recording periods in 2004; what sparks the LP explosion seemingly happening in 2004, why is the LP retained by Valerio for so long as a seemingly very productive pattern alongside the native-like TLP, and why does this TLP struggle to beat out the LP in competition, as it were? It has already been indicated that TLP usage is characterised by what we might call locally recurring MWEs (based on Author subm. a), which means that its use is bound to specific situations. The LP has only one such recurrent expression, as already mentioned. This, as well as the correlation between pattern and function, and pronoun and particle, will be investigated in the present section which takes a closer look at the usage events in the interactional circumstances of Valerio’s biography, as it were, as a learner in this particular classroom.

---

<sup>58</sup> Actually, it fits neither student; if Carlos's data are put to close enough scrutiny, the points in his development at which the patterns were in competition, were not distinctly characterized by one form over the other, either. Valerio’s data, though, are much more elaborate, giving us a bigger material to work with, so the correlations which we are about to investigate are more easily studied in terms of his data.

First, however, it is necessary to document the LP explosion in early 2004. From Table 5 above, it could be deduced that, initially, Valerio used only negation in the first person singular in both LP(s) and TLP. In the recording period immediately preceding the one which marks the above-mentioned explosion, we note one new use of the LP, namely in second person singular. It therefore seems like the use of *you* in negated contexts, at least at this early stage, attracts the use of *no* as the negation particle. This tendency is displayed in Table 6, which shows the type and token occurrences for *you no verb* and *you don't verb*, respectively:

	<i>You no Verb</i>				<i>You don't Verb</i>			
	Token	Type	Ratio	Extract presence	Token	Type	Ratio	Extract presence
Win 04	10	3	0,30	2, 3a-3c, 4	2	2	1,00	7a
Spr 04	9	6	0,66	5-6, 7a	4	4	1,00	7b-d
Sum 04	10	8	0,80		6	3	0,50	8

Table 6: type and token occurrences for *you no verb* and *you don't verb*. Extract presence marks the periods from which the instantiations in the Extracts below are taken.

Even though the absolute numbers in Table 6 are fairly small, it would seem that there are noteworthy statistical differences between the *you + no* and *you + don't* co-occurrences. The *you + no* co-occurrence has been shown to out-compete the TL variety of the pattern at this stage. Another conspicuous aspect of the numbers is the relatively low *you no verb* type-token ratio in Winter 04, the period referred to as marking the LP explosion, which suggest the presence of a recurring MWE. Looking closely, it becomes apparent that this explosion rests on one particular type which recurs 5 times within a fairly dense period of time, namely *you no write*. It is also quite peculiar that Spring 04 sees 100% creativity for TLP use; as will become clear in the next section, however, the numbers cover the fact that three out of those four uses are repetitions of on-going classroom discourse. In order to investigate the reasons for the *you no write*-explosion and the 4/4 creativity of the target language pattern, it is therefore necessary to look in more detail at the classroom affordances (van Lier 1996, 2002) to investigate whether the nature of the classroom interactions may prompt specific usage events and hence learning of specific patterns at certain points in time. This investigative step, to some extent drawing on conversation analytic tools, is necessary to

understand the situated nature of certain points on the general developmental path which in turn help shape the emergent linguistic inventory.

### 5.1 Tracing *you no verb* interactionally.

Extract 1 shows Valerio's first use of the *you no verb* pattern, l. 21. This use is represented in Table 4 above as a 2<sup>nd</sup> person singular LP instantiation in Autumn 2003.

Extract 1 (Sep 23 2003)<sup>59</sup>

- 1 Valerio do you like portland?  
2 ()  
3 Samcha yes I do  
4 ((Valerio writes, eyes fixed on paper))  
5 Valerio and you working?  
6 ()  
7 Samcha yes I do  
8 Valerio ((looks up from paper, makes eye contact with Samcha))  
9 Samcha yes [I am?  
10 Valerio [you working?  
11 Samcha ((squeaky voice, nodding)) ye:s  
12 Valerio wha:[:t ]  
13 Samcha [I am] working  
14 Valerio where  
15 ()  
16 Samcha restaurant [yeah  
17 Valerio [ts a::hr no ((shakes head))  
18 Samcha ((squeaky voice)) what?  
19 Valerio yes?  
20 Samcha ↓o:h  
21 Valer → you [no working ]  
22 Samcha [yeah ((nodding))] ye:s ((squeaky voice))

---

<sup>59</sup> Transcript conventions: () = pause; (...): transcriber's comments; : = prolonged phoneme; underlined item(s): stressed item(s); [ ] = overlap; arrows up / down = rising / markedly falling intonation; → = target utterance; \*...\* = silent voice; U = unidentified student. All names are pseudonyms.

23 Valerio tseh ((looks down, writes))

24 Samcha I give ((points V on shoulder)) you card tomorrow ((gestures)) card [xxx

25 Valerio [uh alright

In this interaction, recorded on September 23 2003 2 months and three weeks into Valerio's ESL classroom career, the students are doing a so-called 'free movement task'. This is fairly common praxis in this classroom, the students moving around freely engaging relatively briefly in cued, serial dyadic interactions (Hellermann and Cole forthc.). As such, it does not really give us any clues as to why this particular pattern is deployed at this particular point in time. Previously, the students, in pairs, have been preparing questions to ask as they move around talking to other peers. The task concerns the practicing of short answers, *yes I do* and *no I don't*. Initially the interaction follows the layout of the task along the lines of questions-answers adjacency pairs. However, Valerio initiates a task expansion, i.e., takes the interaction outside the scope of the teacher's task assignment (Hellermann and Cole forthc.), as he, for some unclear reason, challenges (Koshik 2003) Samcha's task-specific second pair part, l. 7, in which she informs him that she has a job. In l. 8 the interlocutors establish eye contact and Samcha repairs the form of her second pair part answer, changing it from *yes I do* to *yes I am*, thus making her utterance formally align with Valerio's use of the progressive. Samcha seems not to be sure of the relevance of the repair, however, which is try-marked (Sacks and Schegloff 1979, quoted in Brouwer 2003: 540) as indicated by the rising intonation (l. 9). L. 10 marks the first verbalized disbelief by Valerio. Ll. 11 to 14 continue the thread, with Valerio challenging Samcha who in return maintains her point of view. This disagreement culminates, as it were, in the present target utterance in l. 21, *you no working*. The sequence is eventually closed as Valerio accepts Samcha's compliance to bring forth proof of her employment, as it were (ll. 24-25). Valerio expands the task linguistically in l. 10, *you working?*, and maintains his challenge (l. 12 *wha:t?*, l. 14 *where?*, l. 17 *ts a::hr no* and l. 21 *you no working*). Samcha, on her part, insists on her version. Thus, they co-construct the interactional expansion of the task by virtue of what their utterances do, namely initiate and maintain a longer disagreement sequence which is eventually closed down successfully.

Chronologically, the next use of the 2<sup>nd</sup> person negation pattern, TL or LP, found in the data comes in February 2004, a good 4½ months after the interaction extracted above.

Extract 2, Feb 13 04

- 1 Valerio Ramiro  
2 ((the two students approach each other))  
3 Ramiro yes  
4 Valerio are you: get a married  
5 Ramiro no I am married now  
6 Valeiro next year?  
7 Ramiro I am married  
8 (1)  
9 Valerio→ you no married n[ow  
10 Ramiro [yes I'm mar[ried this year ((pointing))  
11 Valerio [you is ((shaking head))  
12 ??? excuse me[: u::::::::::::::::::h]  
13 Valerio [ah heh heh heh] hh

Extract 2, recorded in-between Extracts 3a and 3b below (these interactions take place on the same day) displays a use of a negated string *you no married now* that is parallel with the one encountered in extract 1 above. While it may be postulated that this negation is a learner variety of a target *copula negation* rather than a *do-negation*, this remains speculation. In fact, because of interactional parallels between this extract and Extract 1 above, it makes more sense to think of the two negated constructions as similar because of the similar job they do in those two interactions. Also the task-context is similar, the students doing a free movement task in which they are to ask each other questions and then write down their names if they say yes. Valerio asks Ramiro about his marital status. At first he misunderstands Ramiro who answers *no I am married now* (l. 5), probably orienting to *no*. He then goes on to ask if he is going to marry next year, and Ramiro repeats the message from l. 5 that he is married. A pause follows, indicating trouble, and Valerio then challenges Ramiro in much the same way as was the case in extract 1. Also parallel with the extract above, he maintains the challenge (l. 11) even though the co-participant counters it. The interaction is cut off by another student who wants to do the task at hand with Valerio.

Extract 3a (13-feb 04)

- 1 Valerio ((gets up)) mariela are you () move



2 Mariela no

3 Valerio no? () what? ((turning to Kate))

4 Kate xxx

5 Valerio is all ((motions circling, gets up)) () are you change jobs?

6 Kate no:

7 Valerio → okay () the people [say no you no write name ]

8 Kate [okay xxx ((flipping pages))]

9 Valerio the people say yes I am you write yo the name

10 Kate xxx ((pointing to paper))

11 Valerio you write the name () Roberto yes ((pointing to paper)) y he say yes I am you say

12 [yeam

13 Kate [you: are you change job

14 Valerio uhuh () and Roberto he say yes

15 Kate no [xxx

16 Valerio [no no no () Roberto he say yes you say Roberto ((points to Kate's paper))

17 Kate yes ((nods))

18 Valer → Roberto he say no you no write name ((motions writing, shakes head))

19 Kate ah yeah?

20 Valerio for all the () different ((motioning circling; starts to walk away))

Extract 3b (13-feb 04)

1 U Valerio () are you going to buy a house

2 Valerio → no: ((makes noises)) no: you no write ↓no: you write na:mes ((orienting to paper)) the

3 the people say yes I ↑am you say the ↓name () the people say ↑no you ↓no ((waves))

4 U nothing

5 Valerio → you no write no↓thing

6 U o:h ((pads v on shoulder)) thank you

Extract 3c (13-feb 04)

1 U are you buy how () you buy a how

2 Valerio are you buy a ↑house ((reading partner's question)) no

3 U ↑n[:o

4 Valerio [I have now  
5 ()  
6 U okay () [xxx ((starts writing))  
7 Valerio → [no: this is ((points to partner's paper)) for the names () people say yes I ↑am  
8 you write the name ↓yes alright. no yes I am you write ↓name () you no write no I am,  
you no write yes I am. you say no I ↑am you no write nothing () and you you check  
who people say ↑ye:s and you write name  
11 U ah  
12 Valerio only yes () no I ↓am ((walks away))

Extracts 3a-c above show the learner pattern explosion in its interactional context. In these extracts, Valerio does not challenge his co-participants in the manner of the two previous extracts. He now assists his fellow classmates in getting the task right. This prompts the use of the target pattern, *you no verb*, and its thus locally recurring instantiation, *you no write*, which Valerio uses to instruct his peers on how to perform the task at hand. This applies to all the *you no write* uses above. Note, though, the co-constructed *you no write nothing* in Extract 4b which Valerio deploys on his own in Extract 4c, showing signs of micro-genetic development (Hellermann and Cole forthc.).

Extract 4 09 March 04

1 Ivette David please ((waves David off)) tseh heh heh [hoh  
2 Valeri→ [no you no sa\_ you no write (1) only  
3 ↓she ((pointing repeatedly at Ivette))  
4 Ivette please () [close your hands ((crosses arms; directed at David))  
5 Valerio [you espeak ((moves hand to mouth; directed at David))  
6 David okay ((puts pen away, crosses arms))

In this interaction, recorded approx. one month after extracts 3a-c, Valerio and two classmates are doing a pronunciation tic-tac-toe. Two students are assigned as 'x' and 'o' respectively, and the third student does the checking on a piece of paper on which the words in the game are written so as to represent a game of tic-tac-toe. David is 'x' and Valerio is 'o' and Ivette writes. The pronunciation task lies in the words being phonetically very similar (e.g., minimal pairs such as pen – pan). David is supposed to start the game but keeps breaking the rules, either by pointing to the word he tries to

pronounce or by marking them on the paper with his pen (i.e., doing Ivette's job). At this point in the interaction, Ivette and Valerio increasingly co-instruct David to do the task properly. First Ivette tries to keep David away from the paper physically (l. 1). Then Valerio instructs David not to write, again deploying the learner pattern MWE *you no write* (l. 2), Ivette asks him to keep his arms crossed, presumably to keep him from writing (l. 3), and Valerio finally tells him that he is supposed to speak only (l. 5). David accepts and puts the pen away, and the task gets underway. Valerio's turn-internal self-initiated self-repair (l. 2) may be a trace from previously in the interaction, where Valerio tells David: "you say and she write". If that is the case, then Valerio may be doing a self-repair in l. 5 where he uses the more appropriate verb *speak*.

So far, the *you no Verb*-pattern has been shown to have its first instantiation, *you no working*, as Valerio did a challenge in an expansion of the task at hand. The *you no verb*-pattern explosion, as displayed in extracts 3a-c, takes place in a further task expansion activity, namely one in which Valerio relies heavily on *you no write* to instructs his fellow students. The successive use of *you no write*, displayed in extract 4, also took place in a task instruction environment. Just like the *denegation* pattern is item- and usage-based, its primary exemplar *I don't know*, the learner pattern, *you no verb*, also seems to some extent to be linked in ontogenesis to certain specific items, primarily a locally recurring MWE *you no write*, which Valerio exclusively deploys for the purpose of instructing classmates in the task-at-hand.

As implied in the statistics presented in Tables 4-6 above, the learner pattern evolves productively. In those statistics, there was even evidence that this learner pattern is more productive than the target language pattern, underlining the non-linearity in development towards a linguistic inventory of increasing approximation to target language standards. This productivity of the pattern, found in the statistics, is substantiated locally and concretely in the extract below.

Extract 5 (April 02 2004)

1	Teacher	mhmokay alright so you change that ((reading)) at the time I didn't work for two
2		months okay very good () ((reading)) I was just married for two and a h_ okay I'm so
3		happy with my family I have had one child hi:s name is Andres [xxx
4	Valerio	[mhm his
5	Teacher	he () he what

6 Valerio so very  
7 Teacher he  
8 Valerio he he na his name is Andres  
9 Teacher uhuh  
10 Valerio all time so very happy () and sometimes crying because he don't like mo:m and dad  
11 say no please ((reading)) ((laughs))  
12 Teacher he doesn't like his mom?  
13 Valerio no  
14 Teacher o:[h  
15 Valerio [no [oh  
16 Teacher [now  
17 Valerio ye:[s  
18 Teacher [ye[s  
19 Valerio [he like [en  
20 Teacher [xxx first faze  
21 Valer → he like en but he no like en for me and and my wife she say no Andres please you no  
22 jumping you no oich: ((shaking fingers)) he's wow very very ((waving hands)) and ()  
23 he he have much adrenalina  
24 Teacher ((laughs)) he has a lot of energy  
25 Valerio yeah

In Extract 5, Valerio and the teacher are going over Valerio's written assignment in which he has produced a text about his family. Valerio (l.10) starts elaborating on the part of his assignment which concerns his son. In l. 12 the teacher initiates a repair, the repairable being the part in Valerio's turn where he says that his son does not like his mom and dad (when they tell him no to do things). Valerio eventually does the repair in ll. 17-19. He then tells the teacher how the child is very energetic and does not like it when his mother and father tell him to stop jumping and the like (ll. 21-23). So even though this is reported speech, the use of the *you no verb* carries traces of 'instruction'. For the present purpose, however, this extract is primarily meant to illustrate the productive nature of the learner pattern, as Valerio says *you no oich*; even a nonce item may occupy the slot in the pattern which coerces the meaning (Taylor 1998) of said nonce item to that of the pattern.

From a UBL vantage point, then, it is now possible, given the productive nature of the pattern, to substantiate the claim that the linguistic inventory at this point in development is in fact biased towards the use of the learner pattern rather than the target language pattern for 'instructional' purposes. The claim of productivity of the pattern is also underlined in the following extract, recorded one month after extract 5, in which the pattern is no longer dependent on an 'instructional environment'.

Extract 6 04 May 2004

- 1 Sam arh () I lo:ve chocolate I just [xxx  
2 Valerio → [you no cleaning washing you face in in the night? () you  
3 no [us↓eh () what d'you useh]  
3 Sam [((nods)) xxx ] what?  
5 Valerio what do you useh?  
6 Sam what? [what do xxx  
7 Valerio [what do you use?  
8 Sam clerasil  
9 Valerio yeah ((nodding))  
10 17 lines omitted  
11 Sam xxx but it doesn't help there  
12 Valerio do [↑you\_  
13 Sam [I have to go get the other one. I don't know. probably clerasil is not good  
14 Valerio → nheh heh heh hi::hh (1) no no maybe you no use it correct you no use correct  
15 (1) maybe  
16 Sam maybe I I don't know

In extract 6, there are four instantiations of the pattern, three *you no use* and one *you no cleaning*. Sam and Valerio are doing a task on present perfect, but in this interlude from the task they are talking about the relation between chocolate consumption and acne. Valerio asks Sam if he doesn't wash his face in the evening (ll. 2-3) and then goes on to ask what he uses specifically to get rid of acne. After an other-initiated self-repair sequence, the repairable being Valerio's *what do you use-* utterance, Sam tells him he uses Clearasil, which Valerio seems to be familiar with. Sam then

embarks on an account of Clearasil's insufficiency, most of which is omitted here, which ends in l. 11. Valerio, again deploying the *you no verb*-pattern, then suggests that Sam does not use it correct, which Sam is not sure about. The interlude eventually closes by the participants agreeing that they do not know why Clearasil does not work. The most important thing to note from a language learning perspective is that the pattern under investigation has shown itself to be not only linguistically productive (i.e., deployable with an increasingly wider range of verbs) but also interactionally productive (i.e., portable into new environments).

The productivity of the *you no verb*-pattern as evidenced in the type and token counts (Table 4) cannot be explained in linguistic terms alone or a predetermined acquisitional path from external negation via internal negation to target-language negation as predicted by the developmental sequences. Rather, by looking at Valerio's general developmental tendencies in terms of the classroom requirements as well as the more linguistic side of things, it is possible to infer that at this point in development Valerio interacts with his peers in ways that seem to call upon the deployment of certain linguistic forms which do certain things in certain local contexts, more specifically his determination to assist his peers requires the deployment of a pattern which does the assisting, as it were, namely the instructional (imperative-like) *you no V*. Linguistically, the deployment of this particular pattern at this particular time is a matter of using resources at hand. In February 2004, Valerio was yet to use the target-like *you don't* pattern. The learner pattern, on the other hand, was already in use, as displayed in Extract 1, and thus part and parcel on the linguistic inventory. The fact that the learner pattern, following the usage explosion witnessed in Extracts 3a-c, evolves into such a strong pattern in the competition against the target language pattern, as displayed in the type and token counts, may in large part be due to the successful repeated use of *you no write*, as seen above. From a UBL perspective, with the emergence of *you no write* in recurring interactions, the learner pattern becomes a statistical preference in Valerio's linguistic inventory, its use in later interactions a matter of the frequency-biased (N. Ellis 2002) build of his linguistic inventory – he has not incorporated enough TL patterns into his linguistic inventory in order for it to out-compete the learner pattern.

The more target-like variety of the 2<sup>nd</sup> person *do-negation* pattern, *you don't verb*, is found in the data to occur for the first time in March 2004. Whereas the *you no verb* pattern was found to overwhelmingly occur in instructive environments, the target-language structure equivalent cannot



In this extract, almost two months after Extract 7a, Valerio seems to be recycling Sam's turn. Sam mockingly complains that Valerio is reading out loud. Valerio returns the complaint as Sam was just reading out loud, too. They turn the interaction into a joke. Note how Valerio's turn in l. 5, the target utterance, seems to carry traces of the learner pattern, a distinctly audible 'n'-sound coming between *you* and *don't*.

Extract 7c May 18 04

- 1 Sam choose two (1) two of two of these  
2 (2)  
3 Valerio [but you write\_]  
4 Sam [xxx ] you don't understand ((points)) just ask the teacher  
5 (1)  
6 Valerio you: ((points)) you told me  
7 Sam I don't understand either this ((orients to own book)) (6) I don't understand it  
8 Valerio → what you don't understand?  
9 (3)  
10 Sam I don't understand (1) nail file ((reading))

In extract 7c, two weeks after 7b, Sam and Valerio are doing pair work on vocabulary items. They are to put themselves in the shoes of people who made a time capsule in the 60s and then choose two items from a vocabulary list which they thought people might have put in that time capsule. They are having problems agreeing on what items to choose, partly because they do not understand the meaning of all of them. In l. 8, Valerio again recycles Sam's utterances (l. 4 + l. 7).

Extract 7d May 25 04

- 1 teacher nineteen sixty-four. () plastic heart valve ((over-pronounced)) hm hrm hrm: ((cough))  
2 (1)  
3 Sam what is it.  
4 (3)  
5 teacher well what do you think it [is  
6 unknown [valve  
7 (1)



8 teacher what's a valve a valve is inside your heart

9 multiple xxx

10 Valerio valve uhblɹ ((pointing to chest)) () plastic valve a[:h yeah ((nodding))

11 teacher [it's plastic

12 unknown xxx

13 teacher uh↑uh

14 Valerio oh yeah ((turns around to face Sam))

15 teacher yea:h

16 Valerio you know? ((dir at Sam))

17 teacher it was the first time that you could we could use plastic inside a human body xxx part

18 Valerio xxx ((facing Sam, touching chest))

19 Sam what?

20 Valerio hn: () for the peh

21 unknown *it's good?*

22 multiple [*((laugh))*]

23 teacher [*I guess ↑so I don't ↑know () it was ↑ne ↓w new [medicine*

24 Valerio [*it's new*]

25 Sam I don't know

26 Valeri → you don't know? ((turns around to face Sam))

27 teacher [*transistor radio*]

27 Sam no ((shakes head))

28 teacher contact lenses what are those () what are those () contact lenses

29 Valeri → contact ((puts finger to right eye)) uh you no useh glasses ((motions glasses to eyes))

30 teacher ah yeah [xxx ]

31 Valeri → [you don't useh glasses it's lit\_ very small] yeah

32 teacher you put little pieces of glass in your ↑eyes uhuh

33 unknown con[tact lenses

34 teacher [contact lenses they were brand new

In this interaction, the italics marking interaction that is parallel, as it were, to that of Valerio and Sam, the class continue to work on the time capsule theme, the teacher introducing the students to

items from 1964. In l. 16 Valerio turns around to ask Sam if he understands the current item, *plastic heart valve*, to which he eventually answers *I don't know* (l. 25). Valerio recycles this string, *you don't know?*, in l. 26, and Sam answers *no*, which closes down the interaction. The next item is *contact lenses* which was also a new invention at the time. The teacher asks (l. 28) the class what contact lenses are and Valerio responds by saying that they are not glasses. To do this, he deploys the learner pattern, *you no useh* (l. 29; the spelling 'useh' implies that his pronunciation is marked as he stresses the 'e'). Interestingly, his next turn marks the first spontaneous use of *you don't verb* as captured on tape. So, in this interaction he goes from recycling *you don't know* to repeating an expression which he has used on previous occasions, *you no useh* to producing *you don't useh*. If this is, in fact, the situation in which Valerio 'acquires' or 'picks up' the target-language negation structure it will be noted that it is done in a very smooth fashion in which there is no distinct line between interaction and acquisition. They blend into each other as Valerio interacts with classmate and teacher alike. There is no 'negotiation for meaning' (e.g., Long 1983) involved, there is no focus on form (e.g., Doughty 2001) or explicit noticing (Schmidt 1990) of any part of the input involved (except, perhaps, for 'contact lenses' which does not seem to instantiate a comprehension problem for Valerio) – in short, uptake cannot be used as a metaphor to account for learning (Larsen-Freeman 2004) and the notion of *input* itself seems a derivative concept. Valerio does not seem to operate on linguistic input in order to improve his mastery of structural abilities in his L2; rather, there is conversation all around him which he takes part in in a manner which is reflexive of the social norms as put forward in CA by way of the turn-taking machinery for interaction (Sacks et al. 1974).

The final interaction to be displayed here is the first one in which Valerio seems positively spontaneous in his deployment of the *you don't verb* pattern. Interestingly, the first verb used in that kind of environment is *know*<sup>61</sup> – just as was the case with *I don't verb*.

Extract 8 Jun 25 04

- 1 Eric            I\_ I heard many Mexico er came\_ came here to earn earn enough money and they will  
 2                    [came back Mexico and=  
 3 Valerio        [((nods))

---

<sup>61</sup> There may be one use 30 seconds prior to the one documented below. It is partially inaudible but it sounds like Valerios says *you don't can useh I don't know you no can working you no can talking you no understand and you no* (2) *it's very very difficult*. The part of it that is inaudible is the first part *you don't can useh* which does not make much sense here. Unfortunately, the co-participant's reaction does not reveal anything.

4 Eric =if you spend all the money you came you here again [earn money heh heh  
5 Valerio [yeah but but it's difficult  
6 because uh:m (2) for example my friends they live here single  
7 Eric uhuh ((nods))  
8 Valerio → is group ((gestures; hands coming together)) but is single you know you don't know  
9 what is xxx you don't know what is legal you don't know what is illegal you know (2)  
10 you don't know nothing  
11 Eric yeah

In extract 8 the two students are discussing the difficulties involved in moving to another country. Valerio, in the targeted utterance, makes the point that it is difficult to live alone (as opposed to having an American family which is his own situation) because such a situation makes it harder to learn the ways of everyday life in the new country. Eric, in l. 11, seems to agree, and the students move on to discuss their work situations.

At this stage in development, neither the learner pattern *you no verb* nor the target-like variety *you don't verb* are tightly coupled with a specific environment and the competition, while still favouring the learner pattern, becomes more equal. After Summer 04, Valerio leaves the class to return a year later. For two months during the summer of 2005 very few *you no verb* patterns are caught on tape and it seems like it is gradually losing out against the target-language equivalent. The development traced here has shown that the learner pattern was more productive for a long time. Initially it was dependent on the activity of doing challenges or in interactions which contained what was referred to as an instructive environment, but later use, increasingly varied with respect to main verb deployment, became less attached to specific environments, suggesting its portable nature and productive schematicity. A parallel situation was found to apply to the learning of the target-like variety which initially was dependent on immediately present affordances in the classroom discourse. Interestingly, the learning trajectory of *you don't verb* resembles that put forward in Vygotskian terms of the Zone of Proximal Development (e.g., Lantolf 2005) which suggests that everything in cognitive development is experienced twice, first interpersonally and then intrapersonally, as individuals gradually gain control over their mental capacities. In case of both linguistic patterns investigated, it has been demonstrated that the learning trajectory displayed in the

tables and figures above could only be explained by investigating the changing nature of interaction in the classroom.

## 6. Summary and conclusion.

Summing up, this research has seen how pattern development is item-based and usage-based. The item-based nature could be demonstrated for the various patterns under investigation:

- the general do-negation pattern emerged from *I don't know* to become increasingly varied and productive
- the learner pattern *you no verb* was initially dependent on a local high recurrence of *you no write* in order to take the lead, as it were, in the competition against the target-language variety
- this target-language variety, in turn, was dependent on two things: 1) a previously recurring pattern as *you no useh* evolved into *you don't useh*; and 2) the locally recurring *you don't know* – of course, the already highly frequent *I don't know* a possible psycholinguistic influence.

Even though in the case of *you no verb* and *you don't verb* this paper has investigated patterns that emerged during the time of data collecting, it was impossible to pinpoint a time of uptake. The present data, that is, do not confirm the input/interaction hypothesis (e.g. Long 1983; Mackey 1999; Gass 2003) that linguistic interaction functions as input for the learner who then actively notices (Schmidt 1990) certain gaps in his interlanguage and restructures (McLaughlin 1990) his internal system at the blink of an eye. Rather, what the data suggest is a fundamental coupling of linguistic development and interactional requirements. It is, in other words, futile to keep interaction and learning apart. At the point in time when Valerio finds himself in interactions requiring him to assist his fellow classmates in getting a task right, he uses the linguistic resources readily available to him. In this case, this resource was a non-native-like, lexically specific pattern *you no write*. In terms of the UBL framework and the importance it ascribes to issues of type and token frequencies, these usage events which prompted the locally heavy use of *you no write*, may have laid the foundation for what was to become a seemingly statistical feature of Valerio's linguistic inventory, namely the co-occurrence in negation patterns of *you* and *no* at the cost of a more native-like *do-negation* pattern.

Even though *you no verb* is a non-native pattern, this research has documented both its emergence and its productivity in certain usage events at certain points in Valerio's development, thus presenting it as a pattern which is part and parcel of L2 learning. Its emergence showed item- and usage-based tendencies as it was spawned by a high recurrence of *you no write* and found to be initially dependent on usage events carrying elements of 'instructiveness'. Furthermore, the very productive nature of the pattern alongside the slightly less productive target language pattern suggests that L2 learning, as a process of adopting increasingly native-like patterns, is non-linear at heart and fundamentally unpredictable. The differences presented between the two focal students, and epitomised in the *you no oich*-utterance above, calls for a focus on L2 development as not only item- and usage-based, and locally contextualized, but also user-based; individual differences, once we get a chance to dive deep into longitudinal and interactional L2 data, may turn out to be so outspoken as to be granted a centre-stage position in future studies in L2 development rather than the grand sweep view of development usually presented in quantitative research on developmental sequences (de Bot et al. 2007).

Both the data presented here and in Author and Colleague (2007) support the idea of L2 learning as item-based. In both cases, the MWE *I don't know* instigates the emergence of increasingly abstract patterns which sanction usage expansion of the negation pattern in terms of the inclusion of other lexical items and, perhaps more importantly in terms of level of abstractness, the deployment of past tense expression. The system that is seen to emerge in this fashion is the gradual abstraction of regularities that link expressions as constructions. The present data on Valerio, however, emphasise that individual variation must be taken seriously to avoid falling for the temptation to state sweeping generalities at the cost of certain more finer-grained, locally contextualized developmental insights.

The developmental tendencies found in these data were found to support some of the findings, and refute others, from research in developmental sequences. More importantly, the present research questions the fundamental starting point for defining those sequences, as the data have suggested that learning L2 syntax is not a matter of context-independent rule-learning across linguistic patterns; rather, it is a matter of construction-dimensional exemplar-deduced tendencies that may or may not become schematized as abstract linguistic knowledge in ontogenesis. It is possible to track pattern development in great detail from the concrete item-based starting point of the patterns to the

possible abstraction of regularities that link these patterns as schemas. Such possible abstraction, however, should not be the default starting point for longitudinal L2 learning studies, as research has shown that not all patterns lend themselves easily to abstraction (Author subm. a). In any case, SLA must rid itself of a compartmentalized view of language; i.e., the idea that lexis and grammar are to be kept apart, in research as well as in teaching. L2 learners simply do not learn the two in a manner that justifies keeping them apart; they are intertwined to the brink of being inseparable, and learners acquire them together, not each in its own paradigmatic vacuum. Tomasello (2003) put it nicely when he said for children that they must learn two faces of grammar: smaller elements and larger patterns. Now it seems that it is time for the field of SLA to adopt this insight and investigate its validity for L2 learning in ever more detail.

#### References.

Achard, M. 2007. Usage-based semantics: Meaning and distribution of three French 'breaking' verbs. In Nenonen, M. and S. Niemi (eds.). *Collocations and Idioms 1: Papers from the First Nordic Conference on Syntactic Freezes, Joensuu, May 19-20, 2007*. Studies in Languages, University of Joensuu, Vol. 41. Joensuu: Joensuu University

Atkinson, D. 2002. Toward a sociocognitive approach to second language acquisition. In *The Modern Language Journal* 86/4: 525-545

Author and Colleague 2007. Are recurring multi-word expressions really syntactic freezes? Second language acquisition from the perspective of usage-based linguistics. In Nenonen, M. and S. Niemi (eds.). *Collocations and Idioms 1: Papers from the First Nordic Conference on Syntactic Freezes, Joensuu, May 19-20, 2007*. Studies in Languages, University of Joensuu, Vol. 41. Joensuu: Joensuu University  
Press <http://www.labschool.pdx.edu/research/current/Eskildsen%20&%20Cadierno,%202007.pdf>

Author subm. a: Constructing another language – usage-based linguistics in second language acquisition.

Author subm. b: What's new? – routines and creativity along a usage-based path of second language learning.

Author subm. c: Second language learning as participation and acquisition: towards a new SLA eclecticism.

Bardovi-Harlig, K. 2002. A new starting point? Investigating formulaic use and input in future expression. *Studies in Second Language Acquisition* 24/2: 189-198.

Barlow, M. and S. Kemmer. 2000. Introduction: A usage-based conception of language. In Barlow, M. and S. Kemmer (eds.). *Usage-Based Models of Language*. Stanford: CSLI.

Brouwer, C.E. 2003. Word searches in NNS-NS interaction: Opportunities for language learning? *Modern Language Journal*, 87, 4. 534-545.

De Bot, K., W. Lowie and M. Verspoor. 2007. A Dynamic Systems Theory approach to second language acquisition. *Bilingualism: Language and Cognition* 10/1: 7-21.

Doughty, C. 2001. Cognitive Underpinnings of Focus on Form. In: Robinson, P. (ed.): *Cognition and Second Language Instruction* Cambridge: Cambridge University Press

- Ellis, N. C. 2002. 'Frequency effects in language processing – a review with implications for theories of implicit and explicit language acquisition.' *Studies in Second Language Acquisition* 24/2: 143-188.
- Ellis, N. C. and D. Larsen-Freeman. 2006. 'Language emergence: Implications for applied linguistics – Introduction to the Special Issue.' *Applied Linguistics* 27/4: 558-589.
- Ellis, R. 1985/1996. *Understanding Second Language Acquisition*. Oxford: Oxford University Press.
- Ellis, R. 1994. *The Study of Second Language Acquisition*. Oxford: Oxford University Press .
- Firth, A. & J. Wagner. 1997. On discourse, communication, and (some) fundamental concepts in SLA research. In: *The Modern Language Journal* 81/3: 285-300
- Firth, A. and J. Wagner. 1998. SLA Property: No Trespassing! *The Modern Language Journal* 82/1: 91-94.
- Firth, A. and J. Wagner. 2007. S/FL learning as a social accomplishment: Elaborations on a 'reconceptualised' SLA. *The Modern Language Journal* 91: 800-819.
- Gass, S. M. 2003. Input and Interaction. In *Handbook in Second Language Acquisition*. Catherine J. Doughty and Michael H. Long (eds.) Malden, MA: Blackwell.
- Glahn, E., G. Håkansson, B. Hammarberg, A. Holmen, A. Hvenekilde, & K. Lund. 2001. Processability in Scandinavian Second Language Acquisition. In: *Studies in Second Language Acquisition* 23, 389-416
- Hellermann, J. and E. Cole. forthc. Practices for social interaction in the language learning classroom: Disengagements from dyadic task interaction.
- Håkansson, G. 2005. Similarities and differences in L1 and L2 development. In *Cross-linguistic aspects of processability-theory*. Ed. M. Pienemann. Amsterdam: John Benjamins.
- Kasper, G. 2004. Participant orientation in German conversation-for-learning. In *The Modern Language Journal* 88 (4), 551-567.
- Koshik, I. 2003. Wh-questions used as challenges. in *Discourse Studies* 5/1: 51-77
- Langacker, R. W. 1987. *Foundations of Cognitive Grammar, vol. 1. Theoretical prerequisites*. Stanford: Stanford University Press.
- Langacker, R. W. 1988. A Usage-based model. In *Topics in Cognitive Linguistics*. Ed. B. Rudzka-Ostyn. Amsterdam: John Benjamins.
- Langacker, R. W. 1991 *Concept, Image, Symbol*. (Berlin: Mouton de Gruyter)
- Langacker, R. W. 2000. A Dynamic Usage-Based Model in Barlow and Kemmer (eds.)
- Lantolf, J.P. 2005. Sociocultural Second Language Learning Research: An Exegesis. In: E. Hinkel (ed.): *Handbook of Research in Second language Teaching and Learning*. (Mahwah, NJ: Lawrence Erlbaum)
- Larsen-Freeman, D. 1997. Chaos/Complexity science and second language acquisition. In: *Applied Linguistics* 18/2: 141-165
- Larsen-Freeman, D. 2002. Language Acquisition and Use from a Chaos/Complexity Theory Perspective. In *Language Acquisition and Language Socialization*, Claire Kramsch (ed.), London: Continuum
- Larsen-Freeman, D. 2004. CA for SLA? It All Depends... . In: *The Modern Language Journal*, 88/4: 603-607
- Larsen-Freeman, D. 2006. The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics* 27/4: 590-619.

- Larsen-Freeman, D. & M. Long. 1991. *A Introduction to Second Language Acquisition Research* (London: Longman)
- Long, M. 1983. Native speaker/non-native speaker conversation and negotiation of comprehensible input. *Applied Linguistics* 4/2: 126-141
- Lund, K. 1996. Communicative Function and Language-Specific Structure in Second Language Acquisition – a discussion of natural sequences of acquisition. In: E. Engberg-Pedersen, M. Fortescue, P. Harder, L. Heltoft, L.F. Jakobsen (eds.): *Content, Expression, and Structure. Studies in Danish Functional Grammar* (Amsterdam: John Benjamins)
- Lund, K. 1997. *Lærer alle sprog på samme måde? [does everybody learn Danish in the same way?]* København: Specialpædagogisk forlag
- Mackey, A. 1999. Input, interaction and second language development. In *Studies in Second Language Acquisition*, 21, 557-587
- McLaughlin, B. 1990. Restructuring. *Applied Linguistics* 11/2: 113-128
- Meisel, J., H. Clahsen, & M. Pienemann. 1981. On Determining Developmental Stages in Natural Second Language Acquisition. In: *Studies in Second Language Acquisition*, 3/2: 109-135
- Mondada, L. and S. Pekarek Doehler. 2004. Second language acquisition as situated practice: Task accomplishment in the French second language classroom. In: *The Modern Language Journal*, 88, 4, 501-518
- Pienemann, M. 2003. Language processing capacity. In *The Handbook of Second Language Acquisition*. Ed. C. J. Doughty and M. H. Long. Oxford: Blackwell. 679-714
- Pienemann, M., M. Johnston, and G. Brindley. 1988. Constructing an Acquisition Based Procedure for Second Language Assessment. In: *Studies in Second Language Acquisition*, 10/2: 217-243
- Reder, S. 2005. The "Lab School". *Focus on Basics* 8a. URL: [http://www.ncsall.net/fileadmin/resources/fob/2005/fob\\_8a.pdf](http://www.ncsall.net/fileadmin/resources/fob/2005/fob_8a.pdf)
- Reder, S., K. A. Harris, and K. Setzler. 2003. A Multimedia Adult Learner Corpus, *TESOL Quarterly* 37/3: 546-557.
- Sacks, H., E.A. Schegloff, & G. Jefferson 1974: A Simplest systematics for the organization of turn-taking for conversation. In: *Language*, vol. 50, no. 4: 696-735
- Schegloff, E. A. 1979. The relevance of repair to syntax-for-conversation. In T. Givón (ed.): *Syntax and semantics: discourse and syntax* (vol. 12) New York: Academic Press
- Schegloff, E. A., E. Ochs, and S. A. Thompson. 1996. Introduction. in: Ochs, E., E. A. Schegloff, and S. A. Thompson (eds.): *Interaction and Grammar* Cambridge, Cambridge University Press 1-51
- Schmidt, R. 1990. The role of consciousness in second language learning. in *Applied Linguistics* 11/2: 129-158.
- Taylor, J. R. 1998. Syntactic Constructions as Prototype Categories. In: M. Tomasello (ed.): *The New Psychology of Language* (Mahwah, New Jersey: Lawrence Erlbaum)
- Tomasello, M. 1998. Introduction: A Cognitive-Functional Perspective on Language Structure. In: M. Tomasello (ed.) *The New Psychology of Language* (Mahwah, New Jersey: Lawrence Erlbaum)
- Tomasello, M. 2000. First Steps toward a Usage-Based Theory of Language Acquisition *Cognitive Linguistics* 11/1-2: 61-82.
- Tomasello, M. 2003. *Constructing a Language*. Cambridge: Cambridge University Press.



Tummers, J, K. Heylen, and D. Geeraerts. 2005. Usage-based approaches in cognitive linguistics: A technical state of the art. *Corpus Linguistics and Linguistic Theory* 1-2: 225-261.

Unsworth, S., T. Parodi, A. Sorace, and M.-Y. Scholten. 2006. Introduction. In *Paths of development in L1 and L2 acquisition: in honour of Bonnie D Schwartz*. Ed. S. Unsworth. (Philadelphia, PA: John Benjamins)

Van Lier, Leo. 1996. *Interaction in the Language Curriculum. Awareness, Autonomy and Authenticity*. London: Longman.

Van Lier, L. 2002. An ecological-semiotic perspective on language and linguistics. In: Kramsch, C. (ed.): *Language Acquisition and Language Socialization*. (London: Continuum)

VanPatten, B. and J. Williams. 2007. Introduction: the nature of theories. In *Theories in Second Language Learning. An Introduction*. Mahwah, NJ.: Lawrence Erlbaum

## **Chapter 6:**

### **Portable language experience – a brief outline of UBL.**

#### 6.1. Introduction:

The term 'Usage-Based Linguistics' was used by Langacker (e.g. 1987, 1991, 2000) to describe his own linguistic theory, Cognitive Grammar (CG), as rooted in language users' needs as communicators. By doing this he positioned language knowledge and language acquisition in real usage events. The embedded emphasis on actual language use is what allows Tomasello (e.g. 2003b) to refer to a general framework in linguistic theory as 'cognitive-functional linguistics' or 'usage-based linguistics', hence equalling the two terms. This theoretical framework is radically different from 'traditional' approaches inspired by Chomskyan Formalism. The point of the following section is to capture (some) central issues in both Formalism and UBL, highlighting such differences and the manner in which the latter mindset has informed the present research as displayed in the five research papers.

#### 6.2. UBL: Maximalism and other basic tenets.

UBL is a rather inclusive term, as may be inferred by the brief introduction above. Introductory texts by Langacker (e.g. 1987, 1988, 1991) describe 'a usage-based model' of language knowledge as maximalist, non-reductive and bottom-up. Reflecting the theory-internal tendency to contrast itself with Chomskyan linguistics, which is characterised as minimalist, reductive, and top-down, these features are often repeated by others as defining characteristics of UBL (e.g., Achard 1997; Barlow and Kemmer 2000; Tummers et al. 2005). While the emancipation from Chomskyan thinking is part of the identity of UBL, and therefore deserves mentioning, I will only capture the essence of it here and refer to the works cited above as well as the work of Tomasello (e.g., 1998, 2003a, 2003b) for more detailed discussions<sup>62</sup>. This chapter, then, will briefly delineate the core features of UBL in terms of its maximalist, non-reductive, and bottom-up nature and then go on to outline the specifics of UBL as I have applied it in my five research papers.

UBL is concerned with the semiotic function of language; i.e. the form-meaning pairings which language is seen to consist of. These form-meaning pairings are called symbolic units. All units of a language are symbolic units, i.e. fundamentally and essentially identical, described along a

---

<sup>62</sup> Other linguistic works that discuss the differences between 'functionalism' and 'formalism' include Hopper (1987, 1991); Lakoff (1991); Newmeyer (1991, 1998); van Valin (1991); Croft (1995); Croft and Cruse (2004)

continuum of specificity (from fixed, perhaps idiosyncratic, concrete formulas to abstract schematic templates which in turn sanction the single instantiations) and complexity (from morphemes to full utterances). Language knowledge, in this conception, is seen as a structured inventory of these symbolic units. Indeed, Construction Grammar (e.g. Fillmore et al. 1988; Goldberg 1995, 2003; Croft 2003; Croft and Cruse 2004; Fried and Östman 2005), arguably the single most prominent theory of syntagmatic structure within UBL, sprang from an attempt to capture both linguistic generalizations and idioms and other expressions not traditionally seen to be rule-governed in one all-embracing linguistic model, leading to the idea that all linguistic expressions must be psycholinguistically identical in nature (Croft and Cruse 2004).

Although UBL thus makes no a priori distinctions about core and peripheral aspects grammar, the model does differentiate linguistic material. It does so in terms of schematicity and concreteness – and it does so both in terms of linguistic description and psycholinguistic representation. When describing linguistic material, the notion of *construct* is used to denote the specific instantiation under investigation, and the notion of *construction* is reserved to account for the general schema, to the extent that such a schema is relevant to the linguistic analysis at hand. Psycholinguistically, the maximalistic nature of the model, ensuring co-habitation in the grammar (Langacker 2000, Achard 2007), allows multiple storage in the inventory of both *constructs* and *constructions*.

The term construction, then, usually denotes the most abstract cognitive structures in the linguistic inventory. At the core of UBL, including construction grammar approaches to linguistic structure, is the tenet that such schemas themselves carry meaning (e.g., Fillmore et al. 1988; Goldberg 1995, 2003; Taylor 1998; Langacker 2000; Fried and Östman 2005). As such, they are part and parcel of the structured inventory of form-meaning pairings; they are "[s]chematic templates representing established patterns in the assembly of complex symbolic structures." (Langacker 1991: 16). These constructional schemas range from the fairly simple, such as the schema for 'plural', which consists of two symbolic structures, two form-meaning pairings, namely the noun and the plural morph, to the more complex, such as the 'double-object' construction. This construction is exemplified by such constructs as *he gave her flowers, they baked us a cake, she smiled him her love*. What binds these constructs together as one construction is the shared syntagmatic structure and the meaning with which it is coupled, namely that of 'object transfer'.

Our language knowledge, then, is seen as a structured inventory of this kind of constructions. The maximalistic nature of UBL, i.e., the cohabitation in the grammar hypothesis, allows for the co-existence of the concrete example and the abstract schema. In the following sections, 'plural' and 'double object' are discussed as exemplary of, respectively, the cohabitation in the grammar hypothesis and the construction as meaningful in itself.

Borrowed from Achard (1997), 'plural' exemplified as totally concrete and maximally abstract looks like this:

- car + s = cars
- THING + [morph] = more things

In terms of acquisition and storage, children constructing their L1 (Tomasello 2003a) learn the plural in a usage- and item-based fashion; i.e., on a noun-by-noun basis which is dependent on encountered exemplars in use. They learn the plural form of some nouns in some contexts while not doing so for other nouns in other contexts. This means that ontogenetically, there might be acquisitional phases where plural is neither totally specific nor maximally abstract, but exists in a choice of nouns only, as it were. As the child gradually experiences more and more exemplars of plural, the system develops, acquisition thus inevitably 'grounded' (Barlow and Kemmer 2000) in the child's concrete experience in usage events. This implies that the fully schematic language inventory has its roots in concrete usage; the language that is put to use in the usage events is, of course, concrete and lexically specific, but the emergence of the abstract system, based as it is on abstractions from specific instances, is also linked to this concrete usage event. This means that as the child is in the process of acquiring the language, the specific instantiations of a given potentially abstract representation do not disappear, but are retained by the child for future use alongside the more abstract patterns. This is especially so for frequent items (Tomasello 2003a).

This *cohabitation in the grammar hypothesis* (Langacker 2000; Achard 2007) is related to the rule/list fallacy (Langacker 1987) which is based on the widespread assumption that as general rule learning (system-learning) takes over, specific instantiations (items) disappear from memory and use. Instead, the usage- and item-based trajectory of language learning proposed by UBL implies a non-distinction between the 'item' and the 'system', as opposed to formalist thinking, e.g., 'generative grammar lite' (Tomasello 2003a, on Pinker 1994).

Children's language development is thus based on their experience as language users; in learning they depend on the concrete linguistic material they encounter. It is not that they are never creative, but in UBL language learning is seen as rooted in usage. In Lieven et al. (2003), for example, it is the process of reusing linguistic material that determines the course of language learning in the case of their 2-year-old focal child; 2/3 of the language she uttered was recycled verbatim, whereas the rest consisted of partially recycled chunks, utterance schemas, parallel with collocational frameworks in Renouf and Sinclair's (1991) terms (as explored in ESK1 and ESK2 and to be further discussed in chapter 7). It is one of Tomasello's (2003a) points in this respect that it is important that we describe and analyse the linguistic experience of children learning their L1 in terms of specific words and phrases and not in abstract adult categories à la Chomsky (see also Peters 1983). Doing this, it becomes apparent that children are subject to repetitions of item-frames, that constructions are frequently used in given situations which, essentially, is what allows children to deduce more schematic patterns from usage. Thus, as put forward in Langacker's rejection of the rule/list fallacy, there is no dichotomous relationship between the 'ruly and the unruly' as is seen in the words and rules approach, there exists no meaningless algebraic system.

The double-object construction is often invoked in the UBL literature to exemplify the symbolicity inherent in constructions as such, captured in Taylor's (1998: 195) dictum that "the construction's semantics "coerces" the semantic value of one of its parts, such that the part becomes compatible with the construction's overall meaning". The idea is that the meaning carried by the structure itself imposes upon the language user a certain construal, in the case of the double-object construction that of 'object transfer', so that he or she may understand lexical variations within it such as *Mary sneezed John the football* rather than the more typical *Mary gave John the football*. In terms of prototype theory, e.g., Taylor (1995, 1998), it is argued that the more prototypically used for transfer the verb is, the easier it processes. Therefore, as Tomasello (1998) points out, *Mary smiled John the football* is on the verge of acceptability. The reason why we may understand various constructs problem-free, as long as they adhere to the construction schematics, is that we apply the meaning of the construction per se, basing our understanding of creative constructs on previously experienced, more prototypical instantiations of the construction.

In a Chomskyan setting, just to speculate, the acceptability of the constructions would be a matter of the specifications of the lexical entries – keeping in mind that constructions in this approach are taxonomic artefacts (Chomsky 2000). More narrowly defined, in the above-mentioned cases it would be a matter of constraints imposed by the  $\theta$ -roles outlined in the verb's lexical entry (Chomsky 1986, 1988). *Sneeze*, e.g., is intransitive, so there are no Agent or Goal roles assigned, leaving *Mary sneezed John the football* unacceptable. The same is the case if *sneeze* is substituted with *smile*; they are equally unacceptable – no prototypes allowed<sup>63</sup> – the syntactic construction being a result of the workings of the Projection Principle<sup>64</sup>; a mere artefact. The construction has no semantic significance here, whereas in the UBL tradition it is, like all other items in the inventory, a form-meaning pairing. The Chomskyan approach simply fails to encapsulate that "transitivity is a property of the clause, not of lexical items" (Taylor 1995: 221; see also Goldberg 1995, 1998).

The linguistic model thus described is maximalistic, non-reductive, and bottom-up. It is maximalistic because it allows for co-existence in the grammar of constructs and constructions. It is non-reductive because the linguistic description is not reduced to either 'form' or 'meaning' but attempts to capture both in the term 'construction', and because the schematicity of the construction is not reducible to the sum of the constituents, either structurally or semantically. It was shown in the above, for example, that transitivity is inherent to the construction, not to any of the items in the construction. This means that also in terms of semantic theory do UBL and formalism differ; whereas there is general agreement in formalism that "clause structure is by and large predictable from the semantics of the lexical items" Siewierska (1992: 412), semantics according to UBL is fundamentally non-compositional; the meaning of a string of lexical items is simply not the same as the sum of those items (e.g., Langacker 1991; Taylor 2002). Last, but not least, UBL is bottom-up because the concrete instantiations form the root of schematization, as it were. All extracted generalities stem from specific instantiations in real-time usage events; language knowledge in this sense is a form of portable language experience.

---

<sup>63</sup> Prototype theory, permitting "degrees of membership [makes] it necessary to revise, perhaps even to give up, the modular conception of grammar." (Taylor, 1995: 181-3).

<sup>64</sup> Briefly explained, syntactic structure in this vein is seen as a projection of lexical entry information along the lines specified in  $\theta$ -theory (Chomsky 1986a: 81ff). This includes  $\theta$ -Grid, semantic selection (s-selection), and categorial selection (c-selection).  $\theta$ -Grid is responsible for argument structure, the intra-sentential relationships, of verbs and nouns. For verbs, e.g.,  $\theta$ -grid states the number and kind of arguments needed (including matters of transitivity). It also specifies  $\theta$ -roles, e.g., whether a verb requires an Agent, a Patient, a Goal and so on. S-selection makes sure the above-mentioned roles are played by semantically appropriate lexical items; c-selection determines the grammatical wrapping of complements.

### 6.3. UBL vs. UG: logical problems of language learning?

As opposed to UBL's maximalistic, non-reductive, and bottom-up view of the linguistic inventory stands the Chomskyan subscription to a minimalist, reductive, and top-down model of linguistic representation, or language knowledge. This line of thinking follows from Chomsky's attempts at idealizing speakers, their knowledge, and the communities to which they belong. In terms of capturing the essence of language knowledge, the paradigm aims at a model which is self-contained and rule-based; it consists of a finite number of rules which account for the infinite number of grammatically possible sentences in a language. It is minimalistic because sentences and phrases are only specified in the model with recourse to the most general rules possible. It is reductive because it reduces the object of linguistic inquiry to issues of formal combinability, and because the products of rules are never specified in the grammar; they are merely combinatorial outcomes with constructions seen as epiphenomenal. It is top-down because the focus is on combinability, or computability (Achard 1997), of items by way of the generative syntax, rather than on specific instantiations. Language knowledge is therefore seen consisting of rules – and words that are combined by these rules. This is epitomized in the title of the 1999 book *Words and Rules* by Pinker, a prominent Chomskyan follower.

One of the formulations of the UBL framework and its basic assumptions that I adopt in EC and ESK1 comes from Tummers et al. (2005) who list three principles which unite the various linguistic theories under the UBL heading. Also reflecting the general tendency to emancipate from Chomskyan thinking, these principles include the priority ascribed to language use and the rejection of the competence-performance distinction<sup>65</sup>. The basic Chomskyan distinction between Competence and Performance comes from the strong-lived argument that language users' Performance – seen to be characterised by memory limitations, distractions, errors etc. (Chomsky 1965; Jackendoff 2002) – could neither constitute a valid candidate for serious scientific investigation nor be an ideal realisation of underlying linguistic Competence. Thus, the object of inquiry, language, in what Atkinson (2002) calls the defining moment of 20<sup>th</sup> century linguistics, was removed from its natural context of use and instead the underlying cognitive system of the speaker was emphasised. This mentalistic nature of the Chomskyan movement (the cognitive revolution; Miller 2003) and its deliberate detachment from aspects of usage came with an

---

<sup>65</sup> The third principle mentioned by Tummers et al. is the rejection of the rule-list fallacy. See original text for discussion; see also EC, ESK1, as well as Lanckacker (e.g., 1991).

idealisation of language as a cognitive system, focusing on "an ideal speaker-listener, in a completely homogenous speech-community, who knows its language perfectly..."(Chomsky 1965: 3). In other words, Chomsky's urge to treat language as an object of inquiry in the same vein as what is customary in the natural sciences, and his convinced belief "that a significant notion of "language" as an object of rational inquiry can be developed only on the basis of rather far-reaching abstraction" (Chomsky 1980: 219), has not only set him off towards the embrace of the idealisations mentioned above; rather, the idea of idealisation is of "critical importance if linguistic theory is to be pursued along these lines" (Chomsky 1980: 24). As opposed to the Chomskyan approach to linguistics, UBL views language structure as emerging from concrete usage situations. Thus, UBL also shakes the foundations of the competence-performance distinction as it is believed that language use and language knowledge, social interaction and cognition, individuality and sociality are mutually constitutive. This means that UBL operationalises language knowledge, the structured inventory of symbolic constructions, as "linguistic experience" (Tomasello 2000; N. Ellis 2002). In actuality this means that whenever something has been uttered in interaction it is considered part of the linguistic experience of the person who uttered it. The possible application to SLA of this basic assumption that linguistic knowledge is essentially experiential is explored in the five research papers.

Such an experiential view of language knowledge implies a view of language acquisition as strictly coupled with language in use. In terms of mental processes of acquisition, UBL is affiliated connectionism (e.g., Rumelhart and McClelland 1987; N. Ellis 1998; Gasser and Smith 1998; MacWhinney 2000). Both view language consisting of form-meaning associations – patterns of communication – which people learn as a matter of pattern recognition and remembering. When language learners have experienced enough exemplars of given constructions, regularities are abstracted and schematised knowledge is developing. There is in other words, no lexis-syntax division, no abstract rules governing language. The sharp competence-performance distinction is also given up, with performance, locally applied in language events, viewed as a subset of competence (Barlow and Kemmer 2000), and there is no need to posit a specialized acquisitional module in the brain, such as Chomsky's (1965) Language Acquisition Device/Universal Grammar (LAD/UG<sup>66</sup>), because language *is* logically learnable.

---

<sup>66</sup> LAD and UG are used interchangeably here, cf. Chomsky (1986) who refers to UG as a language acquisition device.



To Chomsky, however, it was necessary to posit the existence of a LAD, "the genetically determined language faculty" (Chomsky 1986: 3), because the abstractions and idealizations he had introduced as fundamental to his epistemology implied that language acquisition was a logical impossibility; it simply could not happen without the intervention of a specialized piece of mental equipment. The primary reason to posit a LAD, then, is an attempt to solve 'Plato's problem': How can we know what we know? (e.g., Chomsky 1980, 1986, 1988). Usually referred to as the *argument from the poverty-of-the-stimulus*, the hypothesis is that children come to know things that they cannot have acquired on the basis of the language that they are exposed to in real life. The Chomskyan dictum in this respect is that this knowledge must come from within the child itself, in theoretical terms as part of UG.

According to Cook (1985: 1), the argument from the poverty-of-the-stimulus is a "typical way into the Chomskyan position". It might be a typical way, but arguably not an apt one, given the increasing evidence that the-poverty-of-the-stimulus-problem is not real (van Valin 1991; Tomasello 1998, 2003a; MacWhinney 2004). Commenting on Chomskyan thinking from the outside, it is evident that the poverty-of-the-stimulus-problem is a theory-internal matter that goes well with the rest of the theory, but as proof in the eyes of the outsider its value is, indeed, limited. The point here, then, is that if one chooses to believe in the reality of the-poverty-of-the-stimulus-problem, then it might be suitable as an introductory way into Chomskyan linguistics (see Tomasello 2003a for a similar viewpoint). Thus, the matter of the basic relevance of UG itself depends on subjective predisposition in terms of Plato's problem. Chomsky (1965: 47; emphasis added) even formulates the objects of linguistic theory as "questions of a *hypothetical* LAD". In other words, if one chooses to go with Chomsky in the formulation of abstract rules of grammar as the underlying concept of language, then one will also agree with the proposals of UG. If, on the other hand, one does not follow Chomsky's formulation of abstract rules etc., then one will not necessarily agree with UG and what it entails. This sentiment is formulated in Dik (1989: 6) as the necessity of nativism due to the dissection of language "from the natural and social environment in which it is used and acquired." It is the purpose of UBL in general, Tomasello (2003a) in particular, to reset this imbalance. This is evident in Tomasello's (2003a: 2-3) discussion of the *continuity assumption* in which he denounces both the poverty-of-the-stimulus assumption and LAD, basically arguing that children are generally powerful learners and that language (knowledge) is much more "child-friendly" than the rule-governed formalist model invites us to think. In other words,

Tomasello shows that we do not need specialised language acquisition equipment of the sort posited by UG theory, and that grammar is not as complex in its abstractness so as to be logically impossible to learn *without* UG (MacWhinney 2004).

UBL, then, denounces Formalism and the idea pertaining to it that we possess algebraic (meaningless) procedures for combining meaningful items. Instead UBL argues that "the essence of language is its symbolic dimension with grammar being derivative", that "linguistic constructions are themselves meaningful linguistic symbols", and that language knowledge consists of "a structured inventory of constructions". (Tomasello 2003a: 5-6) Conceiving language thus precludes the relevance of a hypothesised UG, it does not posit any idealisations along the lines of Chomsky's modularity principles. Instead, knowledge and use are seen as integrated and language is seen as part and parcel of our general cognitive abilities.

In a Usage-Based approach to linguistics and language learning, then, no hypothetical construct in the mind is necessary for language learning; rather, learning words and learning grammar are part of the same overall process of learning to use linguistic symbols. Children are faced with the task of learning "two faces of grammar: smaller elements and larger patterns." (Tomasello 2003a: 41-42). As much language learning as seen from a Usage-Based perspective seems to be a whole → part process, where linguistic elements, be they phonemes, morphemes, words, or phrases, are abstracted from instances of language use, the importance of the segmenting ability becomes evident; at the same time, it may be noted, language use is explicitly emphasised as source of learning (affordances, to use van Lier's (1996) term). In fact, in this Usage-Based perspective, language use is "crucial to the ongoing structuring and operation of the linguistic system" (Barlow and Kemmer 2000). Together, these aspects inspire the point that usage determines both acquisition and structure of language. There is no room for Plato's problem (which is Chomsky's problem!); experiential evidence is rich enough for children to accumulate linguistic constructions ontogenetically – construct their language; i.e., learn it as they get acquainted with it based on general cognitive machinery.

The UBL ideas explored so far of language being a structured inventory of essentially identical constructions whose only differences are a matter of schematicity, and children learning the schematic templates by abstracting away regularities in utterances which they encounter in real-time

experience, coalesce in Tomasello's (2003a) overall acquisitional line of thinking that all constructions are acquired by means of the skills of intention-reading and pattern-finding, that language learning is item-based rather than system-based, and that generally children can apply functional-semantic principles in the acquisition of constructions (i.e. grammar) just as they do in terms of word-learning. The processes of construction learning and word learning are therefore fundamentally identical which renders UG irrelevant, since there is no abstract algebraic system to be learnt. The grammar, being a structured inventory of symbolic structures, emerges in use and acquisition; it is not a system onto itself, working alongside the lexicon, as it is often depicted in the words and rules approach of e.g. Pinker (1999). Therefore, it is also generally assumed in UBL that language use is based on recurrent constructions (e.g., Hopper 1998), and that "the aggregate sum of what speakers do in discourse exhibits recurring patterning beyond what is predicted by rules of grammar" (Du Bois 2003: 49). Interestingly, this quote from Du Bois shows that Pawley and Syder's (1983) 'puzzles for linguistic theory' (as discussed in ESK2 and Chapter 7) were only puzzles in terms of Chomskyan linguistics; a UBL framework is straightforwardly compatible with a view of language which gives prominence to partially fixed, partially open frames along the lines of Pawley and Syder's Lexicalised Sentence Stems (to be discussed in Chapter 7). It is an underlying current in the research as displayed in the five research papers and further elaborated in this and the subsequent chapters, that such an item-and usage-based approach to language learning and linguistic analyses is more apt in terms of capturing the ability to act linguistically in a variety of communicative situations, i.e. usage events – with 'recurrent constructions' or 'recurring patterning' (Formulaic Language, to be outlined in Chapter 7) seen as a major resource in this respect, a major part of the portable language experience of the learner.

At this stage, then, language development according to UBL is captured in terms of lexically specific patterns (e.g., Lieven et al. 1997, 2003; Dabrowska 2000; Israel et al. 2000; Tomasello 2000; 2003a; Dabrowska and Lieven 2005), and the developmental path of children constructing their L1 may be captured as going through the following stages: Holophrases → pivot schemas → item-based constructions → abstract construction with language acquisition being slow, gradual, and piecemeal and not about instantaneous switch setting (Tomasello 2003a). Especially in EC and ESK1, the parallels with this suggested path of learning are clearly stated, the research questions framed around the question of whether or not N. Ellis's (2002: 170) characterisation of the process

of language acquisition as going from "formulas, through low-scope patterns to constructions" would be applicable to L2 learning.

Thus, the specificity of constructional pattern learning is linguistically captured in the idea that learning starts from lexical particulars, with constructional commonalities deduced over time. From the usage perspective, the specificity is captured in the notion of usage event (Barlow and Kemmer 2000; Langacker 2000) or language event (Achard 1997), underlining the experiential conceptualisation of language knowledge in UBL. Usage events are seen as the real-life occurrences of language in action, where the linguistic inventory initially starts to take shape in acquisition. They are thus the fabric that keeps social interaction and language acquisition tied to each other, as the increasing schematicity, seen to materialize as language acquisition progresses along the trajectory laid out above, is a direct result of such specific encounters in which specific lexically filled expressions are put to use.

#### 6.4 UBL for SLA: setting the stage

The focus in my research papers is founded on a blend of these philosophical roots of UBL. In investigating the role of formulaic language – operationalized in the five research papers as recurring multi-word expressions, MWEs, used for a coherent and stable communicative purpose – in L2 acquisition, I have attempted to apply a model of language knowledge which, because of its insistence on giving prominence to specific expressions, lends itself easily to empirical investigation. It allows for taking *any* potentially interesting and/or recurring string of items as its starting point and investigate its acquisition without having to posit its existence outside, either beyond or lagging behind, a current interlanguage system of linguistic rules of the learner. This idea is especially evident in EC and ESK1 where I specifically argue against the traditional conception of formulaic language which I see as being based on a syntactocentric<sup>67</sup> tradition of Chomskyan kinship. I argue that the maximalism of UBL makes for a more suitable framework for capturing the role of MWEs in L2 acquisition. In the five research papers, then, I apply to SLA UBL's concern with the attempt to account for all kinds of usage patterns, i.e., all kinds of linguistic expressions – including idioms, irregular constructions, and MWEs – within one theoretical framework without

---

<sup>67</sup> In ESK1, I changed the notion of syntactocentrism to compartmentalization to acknowledge that researchers working with formulaic language (FL) may not, in fact, subscribe to syntactocentrism. This was done partially to accommodate criticism from a reviewer for *Applied Linguistics*. However, I do maintain the FL is traditionally defined in terms of a compartmentalized view of language knowledge, and that the maximalism of UBL entails a different take on FL. This will be further discussed in chapter 7.

descriptive recourse to such notions as a core and a periphery of the grammar (Tomasello 2003b). All units of grammar are essentially and fundamentally identical (Croft and Cruse 2004).

While the inspiration for the present research thus comes from L1 research and the development of an elaborate construction grammar approach to linguistic structure (Fillmore et al 1988; Goldberg 1995) allowing for a descriptive framework which captures *all* linguistic items within one and the same speaker system, the portability of the framework from these areas to the area of L2 studies is not without its problems. In the SLA literature, the differences between adults learning a L2 and children learning a L1 are often pointed out (e.g., N. Ellis 2002; reiterated in EC). Linguistically, the most striking difference between UBL research in L1 and the present research as displayed in the five research papers here, is that the former investigates language acquisition as part of general cognitive development, whereas the latter attempts to situate linguistic development in a L2 in the social setting in which it invariably takes place. This difference probably has to do with relations to previous research in the respective fields; e.g., from a UBL perspective, L1 learning is seen as lexically specific and therefore logically learnable, hence not requiring specialized acquisition equipment in the form of a LAD. In this, it is evident that part of the UBL agenda is an emancipation from a very dominant formal linguistics, inspired to a great extent by Chomsky. The L2 research setting is different. This will be further discussed in Chapter 8 which is concerned with positioning my 'UBL for SLA' framework in a larger SLA perspective; suffice it for now to say that while truly Chomskyan work in SLA certainly exists, e.g., work centred on investigating the role played by UG in L2 acquisition (recently summarized by White 2007), such work is not dominant in the field at large. The point of applying UBL to SLA is not to emancipate from formalism but to ground L2 development in real usage events; to explore largely uncharted territory in SLA, namely the role of 'context' in 'interlanguage', or L2<sup>68</sup>, development (e.g., Tarone 2000; Kasper 2004).

This attempt to ground interlanguage or L2 development in usage events has been an incipient insight in the course of the present research as presented in the five research papers. The first of these papers, EC, was largely framed on issues of viewing Formulaic Language in a non-syntactocentric light; as such that article was only usage-based to the extent that it adopted the notion of language knowledge as a structured inventory of symbolic units. Problems concerning contextualized use and L2 development were tackled later, especially in ESK1, ESK3, and ESK4,

---

<sup>68</sup> The notion of Interlanguage, while an established construct in SLA research, is not without its terminological problems. I will return to this issue in Chapter 8 on SLA.

which all showed the relevance and validity of applying UBL to SLA as a matter of framing L2 development in terms of lexically specific patterns and their evolution into more schematic patterns as they are deployed in (recurring) usage events.

The UBL trajectory of learning as an interplay between the specific and the schematic has been successfully applied in the research papers EC, ESK1, ESK3, ESK4 with a variety of foci and results. Formulas for the purposes of investigating L2 development were operationalised as recurring multi-word expressions (MWEs) to denote the strictly data-driven point of departure: no a priori formulas were chosen to be investigated. The different patterns under investigation all turned out to be item-based in development, though some of the MWEs, *can*-patterns in ESK1, were found to be rather short-lived, which prompted an operationalization which did not include the criterion of future availability which is sometimes found in the literature (e.g., Tomasello 2003a). This short-lived nature of the MWEs in that study further propelled the interest in looking more closely at the actual usage events in which the patterns under investigation were found to occur, leading to a proposal of more locally contextualized versions of *both* L2 development and the theory of UBL, partly building on Larsen-Freeman's (2006) insight that an atemporal decontextualized competence could never handle the analysis and description of that which is fundamentally dynamic.

Methodologically, the research papers and UBL are bound together by the principles of type and token analyses. Tying in well with the notion of linguistic knowledge as consisting fundamentally of building blocks which can only be differentiated on the basis of the degree to which they can be deduced to be schematic in their underlying representation, this was found to be a productive investigative procedure for getting at the possible productive nature of linguistic patterns. Following Tomasello (2000), the methodological framework is one of conducting linguistic analyses on the basis of samples across and within developmental periods. However, as ESK1 indicated that it was necessary to undertake a more detailed investigation of usage events to describe and analyse the emergent linguistic inventory, it became apparent that a different type of framework would be needed. Even though UBL does ascribe importance to such usage events (Barlow and Kemmer 2000; Lancacker 2000) or language events (Achard 1997), they are not defined in any detail in the UBL literature. They are thought of as a form of locus where linguistic, discursive, social, and cultural choices meet in interaction; are vocalised, to paraphrase Langacker (2000: 9). Such a definition, however, is not sufficient in terms of an operationalization of how to study language as

social activity. While the usage event was sufficient, in ESK1, as a marker for the situations in which *can*-MWEs were put to use by my focal student Carlos, the more detailed interactional analyses undertaken in ESK3 and ESK4 required a more elaborate framework. Hence, I invoked some microanalytic procedures from Conversation Analysis (CA) to investigate the recurrent nature of conversational action sequences and their relationship with the patterns put to use.

This *modus operandi* has resulted in an approach which to some extent combines CA and UBL. This does not mean that CA and UBL are unproblematically inter-compatible; there are fundamental epistemological differences that need to be dealt with. Where UBL is mainly interested in seeking out linguistic units in use and exploring their mental representation / psycholinguistic reality, the object of research in CA is most aptly captured in terms of social order. From a developmental, or learning perspective, this also means that where UBL views language learning as the accumulation of linguistic units in the inventory, CA (insofar as it deals with learning at all, a discussion undertaken in ESK3 and to be revisited in chapter 8) frames learning issues around the notion of graded membership in communities of practice (Lave and Wenger 1991). It became apparent, however, in ESK1 that it would be fruitful to invoke an analytical framework for investigating locally anchored uses of MWEs in more detail than the broad and weak notion of usage event invites for. To meet this challenge, CA's microanalytical tools were applied in ESK3 and ESK4 as operating alongside the more linguistic analysis carried out in the UBL framework. The two are therefore to be seen as complimentary rather than compatible. In any case, it has been necessary to expand on UBL and combine it with other approaches in order to make it satisfactorily applicable to my data and in order to accommodate the usage-based perspective of the approach.

ESK2 took the UBL framework in a different direction, showing that the bulk of language knowledge consists of utterance schemas, formulaic frames (Dabrowska 2000), or Sinclair's (1991) semi-pre-constructed phrases. What separates the UBL conceptualisation of utterance schemas and Sinclair's corpus-linguistics derived phrases, is that the former are schematically related to both other utterance schemas at discrete points in ontogenesis and to instantiations of the same utterance schema at other levels of abstractness at other points in ontogenesis, whereas Sinclair's fixed phrases, i.e., the Idiom Principle, as will be discussed in Chapter 7, are seen as opposite to the open-choice principle which guarantees creativity along the rules laid out by a Formalist notion of

syntax<sup>69</sup>. In other words, the UBL utterance schemas are not fundamentally qualitatively different from the rest of the system; they are not set apart from some combinatorial faculty of language, in the sense that the semi-pre-constructed phrases are seen to opposed creative syntax, which operates on semantically empty principles of combinability in a grand sweep, cross-constructional manner (ESK4).

Whereas Formalism would insist on cross-constructional generalisability from a very early point in children's linguistic development, UBL argues in favour of language acquisition as item-based. Cross-constructional productivity of paradigmatic word classes, primarily the verb, is then taken to be an indication of children operating with abstract linguistic categories (Tomasello 2003a). Such cross-constructional knowledge is only painstakingly slowly built in ontogenesis; it will be shown later in terms of SLA that such broadly sweeping syntactic knowledge has traditionally been viewed as the primary marker of Interlanguage development, but that with the application of UBL to L2 development, SLA research seems to have a tool that is far more empirically convincing in describing and analysing idiosyncratic learning steps towards target language community equilibrium, as L2 learners develop and enhance their portable language experience.

The proposed item-based nature of L1 acquisition has been confirmed to also apply to L2 learning in the research papers, most clearly in EC which laid the foundations for the four successive ones. Here, I choose the term learning (for L2) over acquisition (for L1) to denote the idea that the steps towards linguistic mastery seem to be distinct in a number of ways. This, of course, has often been discussed; aspects of motivation, aptitude, exposure, cognitive maturity etc. are often pointed out as differences involved in learning L1 and L2 (e.g., N. Ellis 2002). However, the data presented in the five research articles also point to the idea that the notion of 'construction' may be too idealized to be applicable to L2 development. While UBL makes for an experiential approach to the study of language learning, in terms of L2 learning it seems pivotal to make the point that the degree of 'constructionality' is a matter of the given construction, or linguistic pattern, under investigation. What I mean by this is that there is no empirical evidence at present to suggest that L2

---

<sup>69</sup> Recently, Hoey (2007) has proposed a revised view of Sinclair's division of linguistic material into the Idiom Principle and the Open Choice Principle. Hoey suggests that the features common to fixed phrases are abstracted by the language users in the course of language use to form a base grammar. In this view, the combinatorial options in the Open Choice principle is determined, not by matters of grammaticality, but by matters of conceptual and communicative feasibility, because the grammar controlling the options is itself determined by use. As such, UBL and Hoey's (2007) lexical priming theory are in alignment (as also pointed out in ESK1).



development, at beginning and intermediate stages, reaches the level of full constructional schematicity. This is what led me, in the research papers, to conduct my basic analyses on the basis of patterns that are part abstract and part concrete; what I refer to as utterance schemas, borrowing a term from Tomasello (2000). The end-point visibility sometimes seen to be implied by the 'acquisition' metaphor (Lantolf 2005) is to be avoided, not only because structural completion seems to be elusive for L2 learning, but because the nature of linguistic portability changes with experience. This emergentist outlook on language learning as a constant and never-ending process, with changing requirements of social interaction at the heart of linguistic experience as it evolves, I argue in ESK1, ESK3, and ESK4, makes for a richer starting point for investigating L2 learning than the structural and developmental stringency entailed by a compartmentalized view of language knowledge and an impoverished view of linguistic interaction as found in traditional SLA.

I will return to a discussion of the field of SLA in chapter 8 which will position my 'UBL for SLA' framework in terms of SLA epistemology. The next chapter takes a closer look at the linguistic phenomenon which initially formed the basic research interest for the research carried out here, namely Formulaic Language.

## Chapter 7:

### Linguistic resources and routines – MWEs and other units in L2 learning

#### 7.1 Introduction:

The purpose of this chapter is to relate the findings in the five research papers to more background theoretical issues concerning the nature of 'formulaic language'. I will give a theoretical impetus for looking for formulas – whatever they are – along the dimensions laid out by Usage-Based Linguistics (UBL); i.e., in a non-syntactocentric, maximalistic vein. During this exposition, parts of the historicity of the five research papers will also become more apparent, their foci and approaches going from being related to a specific interest in the state of formulaic language in SLA, to formulaic language in a UBL sense in SLA, to a more general attempt to instigate a new usage-based, emergentist outlook on SLA. The final steps as conveyed by the chronology of the research papers, towards a fuller application of UBL to SLA as well as an embrace of a locally contextualized UBL for SLA will be discussed in chapter 8 on SLA. The focus of the present chapter is the nature of 'formulaic language'.

#### 7.2 Setting the stage.

Including a wide variety of expressions such as pure idioms (*kick the bucket*), greetings (*how are you?*), collocations (*catch a cold*), and frequently occurring utterances along the lines of *I'm sorry to keep you waiting*, the notion of chunks, or Formulaic Language, is not unequivocally defined in linguistic literature<sup>70</sup>. In fact, no exhaustive list of formulaic expressions exists (Weinert 1995; Wray 2002; van Lancker-Sidtis and Rallon 2004); “[D]ifferent approaches emphasise different aspects of what turns out to be a complex phenomenon” (Butler 2003: 187). Nonetheless, it is assumed from a variety of perspectives that formulaic language abounds in natural language (e.g. Bolinger 1979; Nattinger 1980; Pawley and Syder 1983; Peters 1983; Alexander 1985; Langacker 1987; Bolander 1989, Sinclair 1991; Cowie 1992; Nattinger and DeCarrico 1992; Lewis 1993; Weinert 1995; N. Ellis 1996; Moon 1997; DeCarrico 1998; Fillmore et al. 1998; Hopper 1998; Howarth 1998; Myles et al. 1998; Tomasello 1998; Singleton 1999; Bonk 2000; Schmitt and Carter 2000; Wray and Perkins 2000; Bogaards 2001; Harder 2001; Wood 2002; Wray 2002; Butler 2003; Kecskes 2003; Schmitt 2004; van Lancker-Sidtis and Rallon 2004; Nesselhauf 2005; Bardovi-Harlig 2006). This means that from a theoretical perspective, research into the role of these varying

---

<sup>70</sup> The term *Formulaic Language* comes from the title of Wray's (2002) book.

pieces of formulaic language in Second Language Acquisition research (SLA) is confronted with the challenge of a priori operationalizing formulaic language each time it is under investigation.

Given the fact that there are more than 50 terms covering formulaic language (Wray 2002), not all of them carrying the same connotations terminologically nor describing the same reality for language users, neither communicatively nor cognitively, an even remotely exhaustive a priori definition of formulaic language is probably an illusion. Even thinking about it is bound to throw anybody off their feet; a wide range of descriptive continua would have to be taken into consideration, these being at least semantic, pragmatic, and formal as well as prosodic. Furthermore, given the fact that formulaic language is often, psycholinguistically, seen as a sort of facilitator in terms of fluent speech, interactional and/or conversational aspects would probably also have to be taken into consideration, should such an undertaking ever get under way.

Referenced as the terminological touchstone for diverse areas such as functional grammar (Butler 2003), psycholinguistic SLA research (Schmitt and Carter 2004), and L2 pragmatics (Bardovi-Harlig 2006), one of the most influential working definitions of Formulaic Language is found in Wray's work (e.g., 1999, 2000, 2002). As also pointed out in EC and ESK1, this definition is based on psycholinguistic issues of storage and retrieval; i.e., if a linguistic multi-word item is not generated or analysed by the internal mental grammar at the time of utterance it is considered formulaic. Getting to the core of this definition, this chapter traces terminological development in fields of research concerned with formulaic language and/or other kinds of lexical items and will argue that a definition of formulaic language against the backdrop of syntactically generated language is impossible. Because of an overwhelming tendency to view language as compartmentalized it is also pertinent to investigate the division between syntax and lexis. Therefore, this chapter will explore such terms as lexical item, lexical entry, the idiom principle, the open-choice principle, multi-word items, collocation, lexicalized sentence stems, lexical phrase, and creativity. Leading to a proposed descriptive framework of how to deal with formulaic language in SLA, based on principles of usage-based linguistics as outlined in the previous chapter and the five research papers, this chapter will argue that traditional studies of this feature of language in SLA have lacked insight because of native entrenchment in a compartmentalized view of language knowledge.

### 7.3. On starting points: Linguistic units and L2 vocabulary research.

Before going into details on what is formulaic and what is not, however, it is pertinent to answer the main theoretical questions in terms of the mental lexicon pertaining to the exact nature of notions such as *word* and *lexis*. This chapter therefore turns to take a closer look at theoretical discussions in the research in the acquisition of second language (L2) vocabulary and how the L2 mental lexicon is described. If Formulaic Language truly is at the core of language, it is expected to be a recurring theme, if not a starting point, among researchers working in the fields of lexical knowledge, proficiency, and vocabulary, especially productive aspects of vocabulary and vocabulary development. For the purposes of this chapter, only L2 vocabulary will be discussed. However, it should be noted that this chapter in no way presents an exhaustive overview of research in L2 vocabulary; the present discussion remains purely theoretical and will only concern the nature of the unit in L2 vocabulary research and its relation to formulaic language. No empirical L2 vocabulary research will be dealt with; the point is to explore theoretical frameworks in which most optimally to investigate the role of formulaic language in L2 learning.

Meara (1996) notes that vocabulary acquisition studies no longer constitute a neglected field of applied linguistics. The basic item in such research, however, has lacked uniform operationalization. The *word*<sup>71</sup>, that is, has no unequivocal definition (for discussions, see e.g. Carter 1987; Sinclair and Renouf 1988; Goulden et al. 1990; Bogaards 1996, 2001; Singleton 1999). Discussions of what it means to know a word, including aspects of meaning, form, and frequency, are usually carried out in terms of aspects of single word knowledge (Nation 1990; Melka 1997; Henriksen 1999a; Schmitt 2000). Some researchers, e.g., Bogaards (1996, 2001) who defines a lexical item as being at least one semantic constituent, at least one morphological word, and having stability of meaning, adopt a 'multi-word item' approach to the mental lexicon. N. Ellis's (e.g., 1996, 1997, 2003) views on chunking and sequencing also emphasize the existence of larger chunks of language in the mental lexicon beyond the scope of *words*. Such 'multi-word item' views of the lexico-syntactical area imply the existence of a vast amount of prefabricated units, lexicalised expressions – all instances of Formulaic Language – which are not created by means of syntactic processing. Defining the *word* as more than one constituent would seem to break radically with a compartmentalized view of language, implying consequences for delineating lexis and syntax. Similarly, Read (1997, 2004) notes that accumulating evidence of the importance of various kinds

---

<sup>71</sup> Henceforth, *word* (in italics) is used to refer to the traditional sense of the word as an orthographic unit subsuming grammatical forms. I.e., *run*, *runs*, *running*, *ran* are forms of the same *word*.

of formulaic language in terms of both use and storage present logical problems in delineating the lexicon from other components of language. Also Singleton (1999: 19-20), listing a range of researchers who have empirically confirmed this, points out that L2 vocabulary research supports the idea of a blurred distinction between lexis and syntax.

Such multi-word assumptions about the mental lexicon, the idea that formulaic language is at the heart of linguistic knowledge, originally inspired the questions I initially set out to answer in terms of L2 learning. I wanted to investigate where such formulaic language comes from ontogenetically from a perspective of language as such that does not a priori separate strictly the various dimensions of language. As such the approach as advocated in the five research papers, especially EC and ESK1, is in alignment with the views proposed by Bogaards, Read, and Singleton. As N. Ellis (2002, 2003, 2004; N. Ellis and Larsen-Freeman 2006) has increasingly affiliated himself with principles from Usage-Based Linguistics, the research as presented in all five research papers, but most prominently stated in EC and ESK1, is very much inspired and influenced by his work. Along with N. Ellis (but also in alignment, as noted, with the vocabulary researchers listed above), my original idea was that putting formulaic language at the core of language description, language use, and possibly also language acquisition has implications for how the line is drawn between syntax and lexis. Traditionally, e.g. in Chomsky's Generative Grammar, syntax is the underlying guiding principle of language. On the other hand, an integrated view of the lexico-syntactic domain, with formulaic language at the core, implies that 1) huge amounts of our language consist of prefabricated phrases that are not created from scratch by means of syntactic rules at the time of the utterance, 2) that syntax and lexis are not easily separated, in theory or in practice, and 3) that rules of syntax may not constitute the underlying guiding principle of language.

Those were my theoretical starting points – and, as noted by Singleton (1999; cf. above), much L2 vocabulary research in fact supports those points, especially 1) and 2). Given such important insights, I expected to find vocabulary research to be generally centred on some notion of formulaic language which bridges the gap between syntax and lexis. This, however, is not unequivocally the case. Though Henriksen (1999b) calls for more focus on formulaic language, the focus in Henriksen's own model of lexical development and lexical competence is on single *words* (Henriksen 1999a; Albrechtsen et al. 2008) Consisting of three dimensions of knowledge, the model attempts to describe qualitative aspects of lexical competence, subscribing to the view, shared by

many others (e.g. Read 1988; Wesche and Paribakht 1996) that quantitative studies – often mere word counts – of lexical competence are uninteresting. Henriksen's model consists of two continua of lexical knowledge and one continuum of lexical use. Dimension I is “The Partial-Precise Knowledge Dimension”, implying that language users go from having a vague to a more established sense of a word’s meaning, cf. also Haastrup and Henriksen (1998). Dealing with knowledge of a word’s network relations, paradigmatic and syntagmatic alike, and therefore also collocational profile, Dimension II is concerned with “Depth of Knowledge”. Dimension III is “The Receptive-Productive Dimension” in terms of which it is generally agreed that receptive skills precede productive skills and that all lexical items may potentially be productive (cf. Channell 1988; Henriksen 1999a; Henriksen and Haastrup 2000; Laufer 1998).<sup>72</sup>

While Henriksen's model may potentially comply with a focus on Formulaic Language rather than *words*, there seems to be no general agreement in L2 vocabulary studies that formulaic language is in any way central to the formation of the mental lexicon or pivotal to the learner. Such interests seem to be the prerogative of specialized formulaic language researchers. It is widely acknowledged, however, that the language user must acquire the items of the lexicon, be they multi-word items or not, through exposure to and use of lexical items in a variety of contexts. Beheydt (1987) emphasizes the idea that a given *word* may enter into a variety of combinations and may not mean the same in all contexts. This is echoed by Henriksen (1999a: 308). In alignment with the central tenet of the Idiom Principle, to be explored below, that items tend to appear in the same co-occurrences, this can only be learned through exposure to and use of the language. Summed up by Beheydt (1987: 57), “[t]he learner has not really semantized a new word until he knows its morphological, syntactical, and collocational profile as well as its meaning potential.”

Without reflecting further on the order of acquiring the above-mentioned aspects of semantization, one may note that formulaic language is not given any separate treatment in much modelling of L2

---

<sup>72</sup> Dimension II may be likened to Meara’s (1996) associational networks. Depth of Knowledge is a well-established theoretical construct in SLA vocabulary studies, but sometimes researchers integrate Henriksen's two knowledge continua in one, combining Depth of Knowledge and the Partial-Precise distinction (e.g., Haastrup and Henriksen 2000; Liu and Shaw 2001; Read 1997; Wesche and Paribakht 1996). In terms of production, Wesche and Paribakht (1996) and Melka (1982, 1997) seem to adhere to the idea that the development towards deeper word knowledge of the sort implied by Henriksen’s Dimensions I + II automatically entails productivity, whereas the partial distinction between the two knowledge dimensions and the use dimension is a property Henriksen shares with Laufer (1995, 1998), Laufer and Nation (1999), Liu and Shaw (2001). Doing this allows for a more detailed discussion of psycholinguistic issues such as *automaticity*, *storage*, and *retrieval* (cf. also Henriksen and Haastrup 2000; Schmitt 2000; Bogaards 2001; Hultstijn 2001).

vocabulary acquisition; rather it is implicitly suggested that formulaic language emerges in the language user through *word* learning per se – in the acquisition of a lexical item's collocational profile as specified in the item's entry in the mental lexicon. This seems to be parallel with Chomskyan thinking which emphasises smaller items of lexis by way of The Projection Principle, the basic unit resembling the traditional *word* whose lexical entry information ensures that the syntactic combinatorial machinery works according to the specifics of each lexical item. This means that characteristics of lexical items, indexed in the mental lexicon, determine the structural surroundings in which the lexical items in question can appear. Ultimately, this lexicalist approach implies that language production processes are rooted in the mental lexicon whose items will then dictate the grammatical legitimacy of the utterances in which they occur (Chomsky 1986, 1988; Cook and Newson 1996). Essentially, then, much L2 vocabulary research seems to implicitly, if not explicitly, regard the occurrence of stable multi-word items in the lexicon as evolving substructures of well-developed lexical entries in a well-structured lexicon along the lines of models like the one put forward by Henriksen (1999a). Appel's (1996) account of lexical development supports this as he construes the acquisition of associations and collocational profile as taking place at relatively late stages in lexical learning. At this point, then, with the learning of the collocational profile as part of the semantization process, a picture emerges of the lexicon in L2 vocabulary studies that leaves no room of priority for formulaic language. Rather, this comes out as a minor aspect of language use and knowledge which, in acquisition, arises out of the seeming mess of the lexicon as the lexical entries develop. The same idea is expressed in Bonk (2000), who speaks of collocations as connections between items rather than larger bits of language; to him collocations are a result of pieces of lexis combining according to (apparently generative) rules of syntax. Therefore, although I take an interest in investigating the unit in SLA, as evidenced in the five research papers but most thoroughly discussed in ESK2, the research paradigm perhaps most intuitively appealing, vocabulary studies, does not seem to share my interest in taking formulaic language as the starting point.

In fact, I think the notion of 'starting point' is pivotal in this respect. With the role of 'chunks' in L2 learning as a declared research object – i.e., I did not set out to investigate *words* or *syntax* in L2 learning – I initially envisioned the present research as striking root, as it were, in linguistic theories which do not a priori separate language knowledge into compartments. Those theories were most prominently Cognitive Grammar (Langacker 1987, 1991; Taylor 2002) and Construction Grammar

(Fillmore et al. 1988; Goldberg 1995, 2003; Fried and Östman 2005). As outlined in the previous chapter, the views on language knowledge as propagated in these theories are fundamentally holistic; i.e., they aim to capture linguistic generalities on one level (constructions) and in one system (an inventory of constructions) only. There are no compartments, no core, no periphery – as opposed to Chomskyan linguistics. Such is the state of linguistic synchronic description and analyses in those theories. In terms of developmental issues in ontogenesis, however, I realized that such static, systemic models of language are insufficient as reference point; development is inherently dynamic, systemic models of language are not (Larsen-Freeman 2006). Based on this insight, I make the argument in ESK1 and ESK2 that L2 development should always be studied with reference to specific linguistic items in the data, not with reference to a systemic model of competence which cannot be empirically validated to have psycholinguistic relevance to the L2 learners whose L2 development is under investigation.

In tackling the issues pertaining to the role played by formulaic language in L2 learning I have experienced an incipient recognition that the trajectory of learning as proposed by Usage-Based Linguistics requires a starting point which is different, not only from a perspective which holds language knowledge to be compartmentalized, but also from the full-fledged constructional system that is the structured inventory of symbolic units (Langacker 1987) in mature native speakers of a given language. What was needed was a framework which allows for the slow, piecemeal accumulation of linguistic resources of various kinds and various sizes. Such a framework was accessible in the work of Tomasello (1992, 2000, 2003) and outlined and proposed for SLA in N. Ellis (2002). That proposed learning trajectory from formulas via low-scope patterns to fully schematised constructions, as described in the research papers, requires loyalty towards the one-level view of language knowledge as proposed by cognitive grammar and construction grammar – but it also requires loyalty towards maintaining a level of analysis and description, which allows for the slow emergence of increasingly widely applicable linguistic structures which do not necessarily depend on an increasingly abstract linguistic knowledge, as pointed out in ESK1 and ESK2. At all times must we make recourse to the most concrete level possible, starting with recurring multi-word expressions (MWEs) – i.e., formulaic language.

The focus of the next section is the nature of such formulaic language; as will become apparent, however, this notion is as ill-defined as *the word*.



#### 7.4. Idioms, open choices, and collocations.

This section pays a visit to the findings of some of the early researchers who took an empirical interest in a more 'lexically' oriented approach to language, as opposed to more traditional non-empirical syntactically oriented research. The aim is to show where they differ and where they are similar; this is important theoretical baggage in terms of the overall proposal of this chapter, namely that linguistic theories which do not set up borders between syntax and lexis, and other linguistic compartments, are the most fruitful ones when dealing with formulaic language as briefly outlined above and argued in the five research papers, especially EC and ESK1.

Among the first researchers to manifestly give voice to the importance of formulaic language, Bolinger (1979) memorably said that language production is as much a matter of memory as it is a matter of combinability; as pointed out in ESK1, Bolinger's thoughts on the role of memory implied multiple storage in a manner which today seems to have foreshadowed the advent of Usage-Based Linguistics. A few years after Bolinger's influential essay, Peters (1983) outlined a book-length empirically based framework for research into first language acquisition which took formulaic language as its point of departure. She argued that children more often than not initially acquire and use linguistic sequences that consist of more than one word. This early inventory of linguistic multi-word items then forms the stock from which the children derive the syntactic rules of the language which they are learning. Already in 1983, then, Peters argued in favour of fluidity between syntax and lexis, and echoed Bolinger in arguing in favour of storage redundancy.

Pre-corpus linguistics researchers (Stubbs 2007), Pawley and Syder (1983), in an immensely influential paper, plunge into "two puzzles for linguistic theory" which are "nativelike selection" and "nativelike fluency". Their point of departure in terms of formulaic language is the *lexicalized sentence stem* (LSS) which is "a unit of clause length or longer whose grammatical form is *wholly or largely fixed*; its fixed elements form a standard label for a culturally recognized concept, a term in the language."<sup>73</sup> Aiming at "the creative power of syntactic rules (...) of the Chomskyan approach", Pawley and Syder (1983: 193) "are addressing [the problem] that native speakers do not exercise the creative potential of syntactic rules to anything like their full extent, and that, indeed, if they did so, they would not be accepted as exhibiting nativelike control of the language." Instead, language users rely on recurring chunks of language to guarantee fluency; such chunks are our

---

<sup>73</sup> Pawley and Syder (1983: 191-2); emphasis added for reasons that will become clear.

primary communicative facilitators, and Pawley and Syder (1983) estimate that we each have memorized and stored hundreds of thousands of them. The problem, then, is that if a speaker were to obey the generative rules of syntax he would not only produce faltering speech, he would also more or less indiscriminately produce grammatically correct but pragmatically odd sentences like *your becoming my spouse is greatly desired by me* rather than the more conventional *will you marry me?*. This latter issue, in essence, is the puzzle of nativelike selection; how do speakers know which of the many possibilities their grammar affords them are ‘institutionalised’, i.e. *lexicalised*?

Since Pawley and Syder are explicitly attacking Chomskyan viewpoints it might be worthwhile pausing to consider what he himself might have to say about these matters: In Chomskyan terms, the knowledge underlying nativelike selection would be ascribed to the realm of the vague term *pragmatic competence* whose responsibility it is to put language right, situationally, institutionally, and idiomatically (Chomsky 1980; Cook and Newson 1996). However, it is not a very interesting aspect to Chomsky because, ultimately, it falls outside the scope of the underlying, perfect *linguistic Competence* that is the object of his scientific inquiries. It might therefore be argued, already at this point, that formulaic language is most optimally dealt with when not applying a descriptive apparatus inspired by Chomsky. This will become clearer as this chapter progresses through the disentangling of the theoretical issues in research in formulaic language.

Influenced by Pawley and Syder, Sinclair’s (1991) notion of formulaicity, the Idiom Principle, rests on the idea that words are more likely to co-occur, and re-occur in the same co-occurrences, than to be separated islands, in an Open-Choice principle as he calls it, strung together by complex, cognitively demanding syntactical processing in real-time language use. The Open-Choice principle is what Sinclair refers to as the traditional way of describing, and analysing language, according to which any word can occur in a given sequence as long as rules of grammar are obeyed. Epitomising the Idiom Principle, Sinclair (1991: 110) holds that “a language user has available (...) a large number of “semi-preconstructed phrases” that constitute single choices, even though they might appear to be analysable into segments.” In other words, the Idiom Principle drastically limits the choices actually made by speakers in real-time language use. These chunks of language, the ‘semi-preconstructed phrases’, are not essentially products of syntax, which has implications for language knowledge. While also constituting the theoretical problem of dividing lexis and syntax, in terms of language knowledge, language use, and language acquisition, this, in essence, is the problem aimed

at by Pawley and Syder that speakers do not take full advantage of the ‘creativity’ of the combinatorial possibilities of the recursive rules of Generative Grammar.

According to Sinclair, then, the idiom-principle with its stock of ready-made phrases, including collocational frameworks, i.e., chunk-like pieces of language only partially lexically filled (Renouf and Sinclair 1991), is the overriding principle in natural language use with the Open-Choice principle to resort to when demanded by the communicative situation. Native-like mastery of a language, including native-like selection and native-like fluency, in the spirit of Pawley and Syder, is dependent on formulaic language whereas the more cognitively demanding Open-Choice principle, with its syntactically combinatorial infinite possibilities, comes into play when the speaker has no formulaic language to perform with in a given communicative situation. In psycholinguistic terms, as pointed out in ESK1, this means that the language user operates on two processing mechanisms, roughly one for syntax and one for chunks (e.g. Skehan 1998; Wray 2002; van Lancker-Sidtis and Rallon 2004). Neither Sinclair nor Pawley and Syder, it will be remembered, denounce 'creativity', they merely downplay this aspect as evidence from corpora has led them to believe that humans mostly do not use language creatively in a Chomskyan sense.

The notion of creativity can be problematic in an approach to linguistic knowledge which focuses on issues of routines and recurrences. There is therefore a common tendency, as also noted by Weinert (1995), in work on formulaic language to also deal with pieces of language that are not-so-formulaic, as it were. Both Pawley and Syder (1983) and Sinclair (1991), as well as Nattinger (1980), discuss fully fixed chunks of language and partially fixed, semi-preconstructed ones. In other words, there is in their characterisation of language a cline of creativity, from totally fixed via semi-fixed to totally flexible language. While Sinclair's (1991) Idiom Principle covers both fully and partially fixed routines, and his open-choice principle is straightforwardly compatible with a traditional view of a combinatorial syntax, Pawley and Syder (1983) seem a bit more problematic in this respect. On the one hand they intend to avoid separating language into compartments, as they speak of lexicalization and productivity as matters of degree. On the other hand they speak of inserting yet another compartment consisting of mini-grammars in-between productive rules of grammar and fixed lexical items in the lexicon. Nattinger (1980) tackles the same issue as he suggests that such semi-fixed patterns are somewhere between traditional syntax and lexis. Neither Pawley and Syder (1983) nor Nattinger (1980) quite manage to solve this theoretical issue, probably

because of entrenchment in existing, modular models of language knowledge. To my knowledge, Langacker (1987) is the first to suggest a full linguistic model that does not principally separate linguistic knowledge into compartments. Today, this is a widely accepted assumption in UBL and, indeed, one of the principles that distinguishes UBL from formalism, as outlined in chapter 6. In my view, then, it seems clear that early researchers in formulaic language whose contributions have been sketched here, formulated thoughts about language use and knowledge which, in hindsight, are perfectly compatible with the principles behind Usage-Based Linguistics – perhaps one might even argue that UBL's starting points are directly inspired by them. The view that language knowledge should most profitably be seen as a structured inventory of form-meaning pairings of varying complexity and abstractness seems to be a natural off-shot of the assumptions and findings in early texts on formulaic language such as Bolinger (1979), Nattinger (1980), Pawley and Syder (1983), and Peters (1983). In fact, as also argued in ESK2, the puzzling questions raised by Pawley and Syder (1983) are only puzzles for *generative* linguistic theory – UBL straightforwardly incorporates their implied insights and seeks to explain them as the theory assumes language knowledge and learning to depend on recurring linguistic patterns in recurring usage events, as outlined in chapter 6.

Nattinger (1980, 1988; Nattinger and DeCarrico 1992) attempted to develop a taxonomy of varying pieces of formulaic language for L2 teaching purposes. Underlying this attempt was a recognition of the importance of prefabricated linguistic routines. As mentioned, Nattinger initially struggled to find room for these 'prefabs' in existing linguistic models, and a good ten years later, Nattinger and DeCarrico (1992) only make it halfway there, as it were. While they do speak of multiple storage of items, thus echoing Peters (1983) and Bolinger (1979), they maintain the dichotomous relationship between syntax and lexis, placing 'lexical phrases' – their term for formulaic language – as memorized chunks of language in-between the two compartments. Their book-length treatment of lexical phrases (LPs) is worth spending some time on because it is a landmark contribution to research on formulaic language in SLA and because it attempts a comprehensive taxonomy of formulaic language. Defining the LP primarily with reference to the function it performs in communicative events, as also mentioned in ESK1, they set LPs apart from pure idioms and collocations, i.e., frequently co-occurring words, without pragmatic function. Collocations that have been ascribed pragmatic functions are counted as lexical phrases (they give *how do you do?* as an example of this). As such, one can immediately think of a wide range of speech acts in the tradition

of Searle (1962) and Austin (1969) which are LPs, including greetings, partings, exclamations, warnings, promises, apologies, requests etc. With this long line of potential members in the superordinate LP category, Nattinger and DeCarrico (1992) necessarily go further in their definitions. Applying purely structural criteria such as length and formal fixedness, they present four overlapping categories: Polywords, Institutionalized expressions, Phrasal constraints, and Sentence builders. Polywords are short and fixed, e.g., *for the most part, I'll say, as it were*. Institutionalized expressions are also fixed but may be of sentence length, e.g., *give me a break, have a nice day*). Phrasal constraints vary in length and allow for variation, e.g., *a \_\_\_ ago, see you \_\_\_*). Sentence builders are like phrasal constraints but may be longer and allow for more variation; *a \_\_\_ ago* only caters to items that denote a period of time, whereas a sentence builder such as *not only X, but Y* carries no restrictions.

Nattinger and DeCarrico (1992) go on to define the LPs functionally along the lines of 'markers of social interaction', 'markers of necessary topics', and 'discourse devices', each involving sub-categories. 'Markers of social interaction' are subdivided into 'conversational maintenance' and 'conversational purpose marking'; these include, respectively, *how are you?, I didn't catch your name, by the way* and *do you x?, I think that x, I'd be happy to*. 'Markers of necessary topics', on the other hand, have nothing categorical to do with conversational topics; here is where Nattinger and DeCarrico's pedagogical purpose becomes clear. This category concerns topics with which learners are often confronted, such as *do you speak x? I'm from x, where is x?, I like x* etc. 'Discourse devices' are coherence markers, such as *in spite of x, in other words, and so to speak*. This final category implies a further distinction between macro- and micro-organizers, signalling coherence at discourse level and utterance level, respectively. Each of the formal categories, then, is represented in all functional categories, and for teaching purposes one might group the LPs according to either formal or functional characteristics.

While the work undertaken by Nattinger and DeCarrico (1992) is quite an achievement, their categories are not without problems. For one thing it is problematic that the conversationally defined LPs captured in the framework are described without consistent reference to the interactional job they might do in conversations; e.g., *I didn't catch your name* is put down as a 'summons', but the interactional work that a 'summons' does in conversation is to "provide the summoner with the evidence of the availability or unavailability of a hearer" (Schegloff 1969,

quoted in Ten Have 1999: 16-17); *I didn't catch your name* is arguably not very apt at performing that job. Generally, however, the problem is that, as might be inferred, Nattinger and DeCarrico's (1992) impressive taxonomy is infinite (i.e., one could take the sub-categorisation further than they do) and the weakness is that the ensuing framework obscures the phenomenon it is intended to capture. A similar sentiment is found in Wray (2002) who, as will be discussed in the next section below, attempts to define 'formulaic language' by recourse to only one defining principle of whole-unit storage, rather than listing a whole range of terms each denoting its own formulaic feature.

From a UBL perspective, as outlined in ESK1, the whole-unit storage criterion does not work as sole feature in a definition of formulaic language. Nattinger and DeCarrico's (1992) LPs, however, do not tally well with UBL's maximalistic model either, because they maintain the syntax-lexicon division. Nattinger and DeCarrico's own conclusion is that in order for the LPs to find their place in a linguistic model, more tight definitions are needed. It would seem that UBL goes in a different direction; rather than attempting to describe and taxonomize in ever more detail various pieces of formulaic language, it is necessary to open up the linguistic apparatus to embrace all kinds of linguistic units in one framework and describe them all on one level. In that, UBL and Wray (2002) are allies – but defining principles differ. This will become clear in the next section which is concerned with terminological (and other kinds of) development in formulaic language research.

### 7.5 On the right terms.

Recently, van Lancker-Sidtis and Rallon (2004) noted that the field of formulaic language research is terminologically handicapped, as also mentioned in EC and ESK1. They do not stand alone with this view (e.g. Wray 2002; Butler 2003; Schmitt and Carter 2004), as more than 50 terms denoting formulaicity can be found in the research literature (Wray 2002). It seems that the problem at hand is two-fold: it is one of psycholinguistic aspects of storage, access, and fluency, and it is one of linguistically delineating which words collocate, i.e., belong together, in use (Bonk 2000). Some researchers argue that collocations found in corpora should be accounted for in psycholinguistic terms (Hoey 2007), whereas others, for example Wray (2002) refers to collocations as bordering on her definition of formulaic language. Mirroring Wray's standpoint, Bonk (2000) works on common collocations but will say nothing about issues of storage. This was also apparent in Nattinger and DeCarrico's definition of LPs as they excluded collocations without pragmatic function from counting as formulaic language. It seems that along those lines there is also tendency that work in

applied linguistics and language acquisition, be it L1 or L2, is likely to focus on storage issues and therefore subscribe to psycholinguistic definitions of the phenomenon (e.g., Myles et al. 1998; Schmitt 2004; Bardovi-Harlig 2006), whereas work in corpus linguistics is more likely to focus on linguistic issues of which words may or may not go together (e.g., Sinclair 1991; Stubbs 1995). This latter tradition is further divisible into researchers working in terms of phraseology (i.e., investigations into restrictions, semantic, grammatical or otherwise, on collocations) or frequency (Nesselhauf 2005). For SLA, there is also a branch of research working with the acquisition of L2 collocations (e.g., Biskup 1992; Bahns and Eldaw 1993; Howarth 1998; Gitsaki 1999; Bonk 2000; Nesselhauf 2003, 2005). I am getting a bit ahead of myself here, though; findings in research on formulaic language in SLA is the focus of the next section. In the following I review some of the terminology generally applied in the research on FL.

One recurring notion in the literature is 'collocation' which refers to syntagmatic lexical co-occurrences. The term is applied by Lewis (1993: 93) to "describe the way individual words co-occur with others. (...) Possible two-word combinations vary from the totally unexpected novel – free collocation – to the rigidly institutionalised or ossified form – fixed collocation." According to this definition, collocations reside on a spectrum of fixedness, a collocational continuum (e.g., Cowie 1988; Kjellmer 1991; Bahns and Eldaw 1993; Moon 1997; Howarth 1998; Gitsaki 1999; Wray 2002; Nesselhauf 2003, 2005). Inspired by Firth, this collocational continuum *describes* the way individual words co-occur and makes collocations a vital part of any view of language that takes the lexicon as its starting point. In a similar vein, Beheydt's (1987) term 'collocational profile' is basically all about how lexical items converge in context to give rise to meaningful and native-like communication. In e.g. Henriksen's model, though, this seems to be about single word learning, i.e., lexical entry development along dimension 2, cf. above. Even though Nesselhauf (2005: 12ff), as mentioned above, makes a distinction between frequency-based and phraseological approaches, their descriptive apparatus is the same in terms of the spectrum of fixedness. They differ on other matters, for the present purposes most prominently in terms of whether all lexical connections are collocations, regardless of fixedness. So, apart from maintaining a fixedness continuum it could be argued that there is also a continuum of scholars in terms of how fixed they consider collocations to be; from the most frequency-based who use recurrence (co-occurrence more than once) in a corpus to determine collocational fixedness, to the most phraseologically based who construct their idea of

collocational fixedness on the basis linguistic features rather than probability of occurrence and/or frequency matters.

Two traditional aspects of describing collocations immediately stand out: 1) the notion of co-occurrence, and 2) the fixedness / idiomaticity continuum. Some researchers often assume that formulaic language along those two lines defies rules of syntax, posing problems to the language learner (e.g., Biskup 1992; Bahns and Eldaw 1993; Howarth 1998; Bonk 2000, Nesselhauf 2003, 2005), as also discussed in EC, ESK1, and ESK4. In terms of the scale of fixedness, formulaicity, or idiomaticity a distinction is made in terms of semantic compositionality between idioms that are semantically opaque and collocations that are semantically transparent (e.g., Moon 1997; Gitsaki 1999). This latter aspect is related to the fixedness continuum; if the meaning is opaque, the expression is likely to be fixed, whereas if the meaning is transparent, the expression is likely to be a free combination of words. What is also sometimes deduced is that fixed expressions are also likely to be stored as wholes. In other words, storage-as-a-whole is a matter of linguistic fixedness which is a matter of transparency of meaning. As will become clear over the course of this chapter, the taxonomic principles applied and combined here are untenable as seen from the perspective of UBL.

Wray (2002: 9), operationalizing formulaic language as a sequence of words "which is or appears to be prefabricated; that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar", also questions issues of linguistic fixedness and transparency of meaning as defining criteria. Her term, *formulaic sequences*, cannot be limited by issues of transparency because that would reserve the whole-storage characteristic for pure idioms (and perhaps restricted collocations, on which more below). Her definition is deliberately inclusive as she intends for it to encompass all previous terms used while retaining the option to home in on specific characteristics of given aspects of formulaicity. Her agenda is to set herself apart from previous uses of 'formulaic language' and related terms, because she feels that, quoting Weinert (1995), the many terms covering 'formulaicity' have been wrongly accumulated to mean one and the same. Listing over 50 different terms for what Weinert (1995: 182) claims to be "very much the same phenomenon", among them being *collocations*, *lexical phrases*, and *lexicalised sentence stems* to mention but a few of the ones I have mentioned here, Wray (2002) may have a point that they do not refer to the same phenomenon. However, I do not think this is



what Weinert means; it is quite clear from the outline of Nattinger and DeCarrico's (1992) work above that the multitude of terms applied in the research area of formulaic language do not refer to the same linguistic units and that the researchers<sup>74</sup> are fully aware of this, but it is also quite clear, I believe, that researchers working within this area of interest would claim to be 'allies' of a sort, meaning that they are all referring to the same *overall* linguistic phenomenon, and that they have certain research interests in common. After all, the fact remains that the research is dealing with word combinations that are more or less fixed; the problem is, however, that of all the features used to describe the various word combinations, there seems to be no common denominator. In other words, *some* of the combinations, it would seem, share *some* features, among them being semantic opacity, syntactic irregularity, lexical invariability, institutionalisation, and storage-as-wholes, but no feature is shared by all combinations.

With this in mind, formulaic language can only be categorised *ad nauseam* as in Nattinger and DeCarrico (1992) or by reference to one criterion only as in Wray (2002). Solutions in-between cannot lead to exhaustive definitions; in fact, the position taken here, reflecting the maximalistic nature of the UBL model, is that the phenomenon cannot be exhaustively captured by recourse to something which it is not, because it is indistinguishable from other parts of the linguistic inventory. The UBL perspective on formulaic language is the concern of a later section, however; for now, let us continue with a discussion of the whole-unit storage criterion.

Somewhat oddly, I think that Wray (2002) herself tries to conceptualise an overall formulaic language phenomenon to subsume all other terms when she gives her notion of Formulaic Language status as a super-ordinate term, at the heart of whose definition is storage and retrieval as wholes rather than on-the-spot grammatical generations, as pointed out in EC and ESK1. One consequence of this conception, which Wray is aware of (e.g., Wray 2002: 46-7, 73), is that it cannot handle the notion of collocations, since collocations are placed on the aforementioned continuum of fixedness, some of them totally free (controlled only by grammatical constraints) and some of them totally fixed (pure idioms). Along similar lines, Cowie (1988) distinguished between *formulae* (i.e., situational routines such as *how are you?*) and *composites* (i.e., word combinations such as collocations). In Cowie's terms, *formulae* do not belong to the collocational continuum. Working with linguistic aspects of collocations as they appear in texts and potentially cause learning

---

<sup>74</sup> Other researchers who list sub-terms of formulaic language and discuss their differences include Nattinger (1988), Kjellmer (1991), and Moon (1997).

problems, Howarth (1998) and Bonk (2000) also warn against treating collocations as stored as wholes. This is interesting since much work on formulaic language in SLA deals with the so-called restricted collocations, to be discussed below. While it may ultimately be correct that some collocations hang more tightly together than others, that they are more *entrenched* in usage than others (Langacker 2000; Tomasello 2003b), this is hard to capture in either-or psycholinguistic terms of whole/single unit processing. While totally fixed idioms (*shoot the breeze*) should intuitively be thought of as wholes, linguistically as well as psycholinguistically, for the simple reason that a literal word-by-word interpretation yields a different meaning than the one intended by the use of the idiom, the so-called free combinations (*drink tea*, cf. Nesselhauf (2005)) might also be so frequent, salient, institutionalised, and entrenched that they are, indeed, memorised as chunks.<sup>75</sup> Idiomaticity in itself, then, cannot be considered the ultimate determiner of whether or not something is stored as a whole in memory. If it is, then almost no instances of formulaic language are stored as wholes given the fact that pure idioms are very infrequent (Moon 1997). The linguistic fixedness continuum and a generic notion of whole-unit storage, then, do not seem to be compatible with each other.

Pawley and Syder (1983: 209) entered the same discussion, albeit not in terms of a collocational continuum, as they specified that “[w]hat makes an expression a lexical item, what makes it a part of the speech community’s common dictionary, is, firstly, that the meaning of the expression is not totally predictable from its form, secondly, that it behaves as a minimal unit for certain syntactic purposes, and third, that it is a social institution.” In other words, their lexicalised unit is institutionalised and non-compositional.<sup>76</sup> Similarly, Moon (1997), in her definition of formulaic language, works on three criteria, each gradable, of institutionalisation, fixedness, and non-compositionality. Relying on the idea that if an expression’s meaning is not predictable from its constituents then it counts as a lexical item in its own right, it is the same idea that underlies the collocational fixedness continuum. The question is, can compositionality be put on a continuum alongside the fixedness continuum; is it possible to speak of something as more or less compositional? This question of categorisation will be dealt with below; for now let us look more

---

<sup>75</sup> Note also that *drinking tea*, *drinking beer*, and *drinking water* – all free combinations – might not denote identical acts of drinking. For one thing, we do not always consume those three liquids for the same reasons and the act of drinking them each belong to different social situations, as it were. This questions the idea of ‘free’ substitution which will be further discussed below.

<sup>76</sup> The second criterion, that the unit must act as a syntactically minimal unit is not clearly described, and will not be pursued further here.

closely at Pawley and Syder's own lexicalised sentence stems (LSSs) and see how well they adapt to Pawley and Syder's own definition of what is lexicalised and the generally accepted definition of what is processed and stored as wholes as captured on the collocational fixedness continuum.

Pawley and Syder refer to LSSs as (relatively) well-defined complex lexical items. They include in their own list *How long are you staying?*, *I've never noticed that before*, and *Where did you find it?* to name but a few. From the point of view of compositionality, such LSSs are not idioms in the sense of having a meaning not predictable from their internal structure; rather, like most LSSs they are compositionally literal expressions – "all elements are used in a literal sense" (Nesselhauf 2005: 14). This marks a fundamental problem in the taxonomy as captured in terms of the collocational fixedness continuum; truly literal expressions do not comply with the criterion of non-compositionality. Because they are LSSs of the literal kind they must be considered to be compositional which means that on Pawley and Syder's own terms they cannot be lexical items! Failing to comply with the notion of non-compositionality, LSSs cannot be instances of formulaic language, they cannot be chunks, cannot be stored as wholes, and hence they cannot act as facilitators in terms of processing.

This clearly shows the immense difficulties of adequately describing formulaic language in terms of the categories underlying the collocational fixedness continuum, and it will be argued here that in order to solve this issue it is necessary to abandon the idea that semantic compositionality is the rule. As described in chapter 6, UBL takes non-compositionality as its default starting point precisely because the semantics of an utterance is more than the sum of its parts. Consequently, it is also necessary to abandon the free-restricted collocation continuum which is based on the idea of compositionality. Furthermore, it is necessary to give up on the idea of viewing *words* as having literal senses that may be applied in given free combinations which adhere to compositionality.

Not only is the conception of degree of fixedness/idiomaticity problematic; putting it on a par with the binary categorisation of what is stored as wholes/single units is untenable because it is impossible to combine a binary category of whole/single storage with prototypical categories, which somehow must be at the core of the fixedness/idiomaticity continuum. The theoretical problem is, then, one of categorisation, physical and cognitive alike. For one thing, when taking an approach that divides lexis and grammar in terms of psychological validity, one is forced to also

make a category of phrases, the category in-between the 'list' of *words* on the other hand, and the 'rules' of grammar on the other. This, however, is intrinsically unworkable, since the 'list' consists of semantic items whereas the rules are inherently non-semantic. I believe it is impossible to envision a continuum from meaningful *words* to meaningless rules with 'phrases' somewhere in the middle – phrases that would then be part meaningful and part meaningless. That makes no sense. Essentially, this was also the theoretical problem which Nattinger (1980) and Pawley and Syder (1983) failed to solve as they proposed an insertion of a component in-between lexis and syntax, as discussed above. In order to envision that sort of continuum it is necessary to view *all* aspects of language as meaningful, as is done in UBL's vision of the linguistic inventory of a continuum of symbolic units (Langacker 1987; Radden 1992) which insists on a "uniform representation of all grammatical knowledge" (Croft and Cruse 2004: 255). On this continuum, phrases are not seen as constituting their own homogenous category hovering between syntax and lexis; rather some instances of formulaic language would be characterised as idioms (*kick the bucket*), but all phrases would be symbolic constructions; more on this below.

*Semantic opacity* and *lexical substitutability / commutability* are two variables often applied to determine the degree of fixedness and idiomaticity (Nesselhauf 2005). Failing to realise that single words are rarely relevant as a unit for semantic analysis (Stubbs 1995), and that words as a rule are inherently polysemous, (Firth, quoted in Poulsen 2005: 42), researchers in Phraseology base those notions on an explicitly compositional view of semantics, the former referring to the extent to which the semantics of a given item in a formula deviates from its literal meaning, the latter referring to restrictions on the potential substitution of elements in a word combination. If word meanings are literal, substitution is likely to be free, and the combination is not a collocation. If, on the other hand, word meanings are not literal, substitution is likely to be arbitrarily limited and we have a restricted collocation. The third option is that nothing in a given combination is substitutable without radical alteration in meaning and we have an idiom. So far so good. Keeping in mind that phraseology speaks of graded fixedness and degree of idiomaticity along a continuum, it is necessary to ask the following questions in terms of categorisation: can word meaning be more or less literal, substitution more or less free, combinations more or less idiomatic? These must be the

basic principles in constructing a fixedness continuum; without that sort of fluidity there is no continuum, only binary categories where membership can never be ambiguous.<sup>77</sup>

Applied to three combinations which respectively represent free combination, restricted collocation, and idiom and which all include the verb *to shoot*:<sup>78</sup> 1) Shoot a gun (free); 2) Shoot a film (restr.); 3) Shoot the breeze (idm.), it is inferred that semantic opacity and lexical substitutability vary among the combinations. In 1) meaning is literal, semantics fully compositional, the items freely substitutable; i.e. there are lots of things you may shoot, there are lots of things you may do with a gun, without *shoot* and *gun* changing meaning. While I agree that meaning is literal in the sense that it is not metaphorical, meaning, it is argued here, is not compositional, and therefore the items not entirely freely substitutable, unless singular dictionary definitions are the starting point. Of course, in a free combination, substituting an item in the compound should not result in any meaning change in the other element. This, however, is not as clear cut as it may sound. In terms of conceptual semantics this is evident if *gun* is substituted for *arrow* which changes the meaning of the verb. Shooting guns and arrows, or machine guns and rifles, are simply not identical acts. It becomes more troublesome outside that domain, e.g. in the case of *shoot the ball*, *the puck*, etc. or *shoot the bullet*, *the missile* or *shoot a goal*. Are they not representatives of a literal sense of *shoot*? Are they whole new free combinations? Or would they be characterised as somewhat restricted combinations because the verb, in at least the two latter cases, probably does not combine with as many nouns as does *shoot [weapon]*? Are they metaphorical? Are they even identical in terms of the meaning of *shoot*? Of course in terms of *shoot the man*, *the elephant* etc., the meaning of the verb is different. But where does *shoot a hole* (in the wall) or *shoot a syringe* (into the arm) fit the picture? It goes for all these examples that the meaning of *shoot* changes, so meaning is not compositional, the noun is not freely substitutable without affecting the meaning of the verb – and if the meaning of *shoot* is literal in 1) it cannot be literal in the examples with the substituted nouns unless it is agreed that there are as many literal meanings as there are 'free' combinations.

Thus, a word's literal meaning is not easily deduced, and it seems that even in the so-called free combinations the items cannot be freely substituted without meaning alterations. When something

---

<sup>77</sup> Please note that some researchers, according to Nesselhauf (2005), only operate with the substitutability of verb (in Verb Object collocations) based on the idea that the noun is semantically autonomous, i.e., not dependent on any specific verb to materialise its meaning. This has no bearing on the present treatment of the notion of substitutability, however.

<sup>78</sup> Example borrowed from Revier (2005)

in a combination is substituted with something else it changes the whole combination; potential change is not limited to any particular item but to the whole construction. This dictum should be taken seriously by researchers claiming to study – combinations. Instead, failing to take non-compositionality of semantics as the point of departure, most research dealing with collocations, it seems, are studying single words and their combinatorial possibilities as seen from a projection principle viewpoint.

At this point, then, *semantic opacity* and *lexical substitutability* are dismissed as valid variables in terms of determining degree of fixedness and idiomaticity. Nesselhauf (2005), acknowledging the problems in measuring opacity, resolves to apply only substitutability (Nesselhauf's term is 'commutability') in discriminating between free combinations and restricted collocations as well as restricted combinations and idioms. However, since these variables as well as the fixedness continuum itself are centrepieces in traditional research on collocations, and with the impossibility of envisioning a fluid continuum from what is literal to what is idiomatic, the whole 'collocational continuum paradigm' seems untenable to me.

In terms of the restrictions thought to apply to restricted collocations, arbitrariness is a deciding principle. Nesselhauf (2005: 31) discusses principles of semantics in distinguishing between free combinations and restricted collocations “on the basis of whether the nouns that a given verb in a given sense allows constitute a fairly large and consistent [semantic] group or not.” If that is the case, restriction is not arbitrary and the combinations free. If, on the other hand, it is impossible to envision such a homogenous group of nouns, restriction is arbitrary. It is problematic that the restriction lies in the idea that the meaning of the verb determines the substitutability of the object – this is only half the story. Unless the object also allows combination with the verb, there is no collocation. The whole utterance, the construction, does not live its own life without consideration of the constituents – and vice versa. In other words, combinability is a two-way street; one item does not arbitrarily choose the next – or the former as Nesselhauf (2005) argues is the case for *do a favour* where *favour* (the 'keyword') is selected semantically and *do* (the 'value') is chosen arbitrarily by favour (as opposed to another verb, say, *perform*). If this were the case, it would be impossible to speak of formulaic language in the first place because everything (apart from traditional idioms) would be compositionally combined. And this is exactly the consequence of borrowing semantic and grammatical theory from a compartmentalized view of language knowledge, according to

which semantics is compositional and grammar consists of combinatorial rules. The insight to apply, of course, is that *do a favour* and *shoot the mountains* (or whatever) are chosen as constructional units by the speaker because of what they mean and the role they might play in the given communicative situation. Because there is a *shoot X* construction everything that is conceptually feasible and communicatively required might go into the X position. Because we know how to do somebody a favour, the construction may be extended to allow us to do them other 'things' as well, as long as there is conceptual agreement between items in the construction and between the construction itself and the items. *Doing somebody a pleasure* or *performing a favour* for instance, are thus not inconceivable. At all times problem-free deployment of such rare or idiosyncratic expressions it is a matter of the context and the reaction of the co-participants in the interactional encounter.

In terms of the data that I present in the five research papers, it is also evident that the collocational continuum taxonomy is irrelevant for my research purposes. Most of the 'word combinations' that I deal with and that my focal students seem to base their linguistic interactions on, cannot be captured in terms of the collocational continuum of idiomaticity, fixedness, and freedom of combinability. The data-driven and longitudinal nature of my research has required a different kind of linguistic framework for analyses. It transpired that pure idioms or other a priori easily defined 'formulas' or collocations were hard to come by in the data, as my students seemed to be operating on different kinds of units in their classroom interactions. Such units, though some of them did seem to have the psycholinguistic qualities of formulas, i.e., *I don't know*, were more often equivalent to Pawley and Syder's LSSs or Renouf and Sinclair's 'collocational profiles' in that they seemed to consist of part abstract, part concrete bits of language. I chose to operationalise them as 'utterance schemas' to show the theoretical indebtedness to Michael Tomasello and to underline the idea that some form of schematicity was thought to underlie the open slots in the patterns. Such resources are accumulated over time in the L2 classroom, it seems, as learners encounter new interactional settings. In terms of the present research, this is especially evident in ESK1, ESK3, and ESK4.

In other words, it turned out that any strict sense of L2 learning as being based on linguistic chunks that are then analysed to become integrated in the language grammar failed as the default way of looking at L2 development. It also turned out that looking at it the other way around, with the acquisition of 'words' and 'rules' as separate entities, was futile. Both accounts of L2 ontogenesis are

probably too rigorous and too rigid to account for something that turns out to be as complex and chaotic as L2 learning. Rather, L2 development seems to be characterized by the accrual of an assortment of different kinds of linguistic routines and resources, some of them recurring MWEs, but most of them recurring utterance schemas. Such a view of development, it became apparent, is more in tune with a view of linguistic knowledge and language learning which gives prominence to recurring concrete items and which does not a priori distinguish between 'words' and 'rules'. UBL is that sort of framework, with its view of language as permeated with meaning at all levels, stating that constructions are form-meaning pairings along a continuum of symbolic constructions: "grammatical units are meaningful and [...] the differences between lexicon, morphology and syntax are only a matter of degree along a continuum of symbolic units" (Radden 1992: 531).

To illustrate the difference between the two continua presented here, the collocational continuum and UBL's 'schematicity continuum', one of Wray's (2002) own examples will do. She states that *keep your hair on* is only formulaic when it means 'calm down' and not when it means 'don't remove your wig'.<sup>79</sup> This displays Wray's preference for operating with a dual system approach, seemingly compatible with Skehan's dual-processing views, rendering the expression a lexical item in the former sense and a product of the generative grammar in the latter. As is typically done, she takes full semantic analysability and compositionality as her starting point, as humans' default mode of semantic interpretation, thus failing to acknowledge that the reason why we understand the latter is not that we are able to combine our way through the utterance compositionally; we understand it because of our ability to construe the situation in a meaningful manner on the basis of a 'collocational frame' in the sense of Sinclair and Renouf (1991), which, in this case, might be termed the 'keep-construction'. It is a semantic frame with instantiations such as *keep your arms inside, keep your head, voice down, keep your shoes/jacket/etc. on*, and therefore partially formulaic. From this perspective, the *keep your hair ons* are, respectively, an idiom with a rather specialised meaning and use, and an instantiation of a more generally applied 'symbolic construction'. Both are instantiations of the same overall 'constructional schema' (see detailed discussions in Fillmore et al. 1998; Croft 2001; Croft and Cruse 2004). Neither of the examples discussed so far should be seen as compositional combinatorial outcomes of a generative system of meaningless grammar. But everything that does not count as a pure idiom is viewed this way by

---

<sup>79</sup> Given the fact that it is possible to have extra hair, not in the form of a wig, but in the form of extensions attached to one's natural hair, the utterance in fact has a third meaning



proponents of the collocational continuum framework. This approach is not tenable. A change in directions is needed in terms of theoretical point of departure.

Wray (2002), however, states that it would be premature to adopt a descriptive framework such as Construction Grammar (Goldberg 1995, 1998; Croft 2001) because it allows for formulaicity. Because this theory of grammar, belonging to Usage-Based Linguistics, explicitly acknowledges the fact that formulaic language is a widespread phenomenon in language use (cf. Tomasello 2003a), it would not be able to challenge the core assumptions of formulaicity, so Wray (2002: 10) argues. However, it could be argued that Wray, on her part, overlooks the futility in accepting the validity of Chomskyan generative creativity while aiming to take formulaicity seriously. From such a perspective, formulaicity will always be considered deviant, "inferior speech", cf. Van Lancker-Sidtis and Rallon (2004), in submission to the Holy Spirit of a Generative Grammar or some other faculty capable of generating the correct sentences of a language (and only those). Furthermore, UBL appreciates the co-existence of formulaic and schematic language; however, precisely because it does not dichotomise the two, the notion of formulaic language, as operationalised by Wray among others, is deflated. UBL does not need to challenge the core assumptions of formulaicity – it simply does away with them because the fundamental approach is essentially different, with no demarcation line between opaqueness and transparency, idiomaticity and non-idiomaticity, and syntax and lexis. Rather, the starting point in UBL is that all linguistic units are essentially and fundamentally identical, the only difference among them being a matter of schematicity, as outlined in chapter 6 on UBL. A fundamental tenet in my approach to researching L2 learning, this is also explored in the five research papers.

#### 7.6 Formulaic language – findings in SLA

This brings me to a brief discussion of the status of formulaic language in the context of SLA. As pointed out in ESK1, research into the role of formulaic language in L2 learning has not yielded systematic results (Schmitt and Carter 2004). In keeping with a compartmentalized view of language knowledge, however, there seems to be consensus that formulaic language deviates from the learner's combinatorial interlanguage system; either formulaic language is seen to be beyond the learner's interlanguage (e.g., Bolander 1989; Nattinger and DeCarrico 1992; Myles et al. 1999; Bardovi-Harlig 2002) or lag behind it (e.g., Irujo; Nesselhauf 2005). Mastery of formulaic language is traditionally seen as something that operates alongside traditional rules of grammar, therefore

posing problems to the language learner (e.g., Howarth 1998; Wray 2002; Nesselhauf 2005), the tenet being that collocations are an arbitrary gap in an otherwise flexible combinatorial system (Howarth 1998). Thus, some scholars have argued that formulaic language is a performance phenomenon with no links to an independently developing grammar (e.g., Krashen and Scarcella 1978; Bohn 1986; Granger 1998), whereas others agree that formulaic language is developmentally significant and feeds into the rest of the system in development (e.g. Hakuta 1974; Bolander 1989; Nattinger and DeCarrico 1992; Weinert 1995; Myles et al. 1998; 1999; Schmitt and Carter 2000; Bardovi-Harlig 2002; Wood 2002).

It was mentioned at the beginning of this chapter that some of the early researchers in the field of formulaic language may have foreshadowed the advent of UBL. Similarly, some have more recently advocated a view of language knowledge as a formulaic-creative continuum (e.g. Nattinger and DeCarrico 1992; Weinert 1995; Schmitt and Carter 2000). As pointed out in ESK1, the maximalistic model of language knowledge envisioned by UBL entails precisely such a view. The UBL framework with its perceived fluidity among linguistic patterns and its default assumption that language learning is fundamentally a matter of abstracting generalities from recurring utterances, is particularly suitable for capturing the dynamic interplay between formulas and creativity (Myles et al. 1998, 1999) and accounting for the gradual evolution of formulas into increasingly more productive structures. This sits very well with also Bolander's (1989) findings that creative language seems to grow out of formulas.

Viewing language learning as a gradual evolution of recurring chunks of language into increasingly more productive structures, the item-based nature of the UBL trajectory of learning makes for a starting point which assumes that formulaic language plays a pivotal role for language learning. However, it does so without specifically marking formulaic language as such as opposed to language that is non-formulaic. Therefore, in the five research papers, I operationalise formulaic language as *recurring multi-word expressions used for a relatively coherent and stable communicative purpose* (the aspect of stability was added in ESK1 as the data seemed to suggest that the MWEs were somehow coupled with specific interactional situations, thereby retaining their function over time). The key to understanding this difference in approaches to the study of the role of formulaic language lies in accepting the different theoretical starting points (i.e., that UBL is a non-compartmentalized model of language which does not view formulaic language as 'non-

generated') and hence accepting that an investigation of L2 learning as a gradual loosening of chunks, as it were, implies adopting an entire model of language knowledge.

Abolishing the distinction between a core and a peripheral grammar, UBL thus challenges the widespread practice in SLA of viewing formulaic language against a background of compartmentalizing linguistic knowledge, as pointed out in EC and ESK1. Therefore, UBL dismisses the view of formulaic language as being somehow beyond the linguistic capabilities of the L2 learner, an otherwise established feature of formulaic language recurring in the SLA literature (e.g. Weinert 1995; Myles et al. 1998; Bardovi-Harlig 2002). UBL contests all tendencies to essentially differentiate between formulaic language and syntactically driven language resources (Bardovi-Harlig 2006), the former handled by the lexicon and the latter by the grammatical system, as is done in psycholinguistic 'dual processing' models of language (e.g., Skehan 1998; Wray 2002; van Lancker-Sidtis and Rallon 2004). Instead, characterising language knowledge as cognitive templates in a structured inventory of symbolic units (i.e., form-meaning pairings) along a continuum of linguistic expressions ranging from the totally specific to the maximally general in terms of constructional complexity and schematicity, UBL insists on the view of all linguistic representations as cognitive routines emergent in usage events rather than 'stored' linguistic assemblages along the lines of either words or rules (Langacker 1987, 2000; Barlow and Kemmer 2000).

Even though, as pointed out in EC and ESK1, I thus view formulaic language in an entirely different view than is usually done, studies in collocational knowledge in SLA (e.g. Biskup 1992; Bahns and Eldaw 1993; Appel 1996; Granger 1998; Howarth 1998; Gitsaki 1999; Nesselhauf 2003, 2005) as well as recent psycholinguistic SLA research in formulaic language (Schmitt 2004) deserve mentioning here. Starting with the former, I follow Nesselhauf (2005: 8; also briefly reviewed in EC) and note that previous studies in this tradition has reached the following conclusions:

- Collocational production presents problems for learners, and more serious problems than vocabulary use.
- Learners use fewer collocations than native speakers.
- Learners not aware of restrictions (restricted collocations).
- Learners not aware of combinatorial potential of lexical items.

- Results vary in terms of relation to proficiency.

The findings are less than surprising. It is, after all, a truism in SLA that learners rarely achieve native-like proficiency (Selinker 1972); it is not to be expected that an area referred to as 'collocational' or 'phraseological' competence should present itself otherwise. It is pointed out in EC that Nesselhauf's (2005: 69) results do not really imply severe difficulties in the area of collocations. I will just briefly expand on that here. Nesselhauf's data come from a written corpus of 318 essays by 207 different German learners of L2 English in university. From that corpus she extracted more than 2000 restricted Verb-Noun collocations. In order to determine the correctness and appropriateness of the collocations, she used four different dictionaries (see Nesselhauf 2005: 49-50) and she considered their contextual use in the essays. When the dictionaries were ambiguous as a measuring tool, Nesselhauf used four native judges whose average judgement on a five-point scale, ranging from clearly acceptable to clearly unacceptable, was used to determine the satisfactoriness of the collocations. The overall result was that 66% of all collocational usage was clearly or largely acceptable and 25% clearly or largely unacceptable. The missing 9% could not be determined; these collocations were judged to be 'questionable'. Given the methodology of averaging the results of 4 native judges, however, sometimes some of the collocations that ended up as unacceptable had in fact been accepted by one or two of the judges – and vice versa. This seems to constitute a serious reliability issue; what is more interesting for the present purposes, however, is that 66%, perhaps as much as 75%, of all collocational usage was judged to be acceptable by native speakers or dictionary. While it may be argued, on the basis of the numbers alone, that the learners are doing pretty well in the area of collocations, such quantitative methodology as applied by Nesselhauf is dubious in itself. More qualitative analysis of the data is needed, such as inclusion of situation of usage, what sort of collocations are used wrongly, what sort correctly, what are the sources of error, types of error etc.

A recent volume on various aspects of formulaic language in L2 learners of English (Schmitt 2004) presented a range of different, mostly psycholinguistic, studies. The studies from that volume which investigated acquisition and processing of formulaic language by L2 learners will be briefly reviewed here. Some of the studies investigated the acquisition of formulaic language by advanced learners of English as L2 (university students in Nottingham) and found that they enhanced their knowledge of formulaic language during the time of their course at the university, but the research

did not investigate whether it happened incidentally or by way of formal instruction (Schmitt et al. 2004). That eventually became the research question in subsequent studies (Dörnyei et al. 2004; Adolphs and Durow 2004) which focused on 7, respectively 2, of the original 94 informants. These studies found that those who had the most optimal sociocultural integration into the surrounding community were the most successful in learning new formulaic language, suggesting that incidental learning was pivotal.

In a series of studies, Schmitt et al. (2004), Underwood et al. (2004), and Schmitt and Underwood (2004) focused on processing differences between native and non-native speakers of English in controlled laboratory environments, using dictation tasks, eye tracking methodology, and self-paced reading. Native speakers did better than non-native speakers in all tasks, but some of the word-clusters (authors' term) tested in the dictation task were seemingly not processed as wholes by either native or non-natives. The authors offer as explanation issues of transparency of meaning, length of cluster or boredom with the dictation; there was a tendency that the short clusters as well as the clusters presented to the informants at the beginning of the test, as opposed to those presented towards the end, were remembered holistically. The transparency of meaning issue, however, is probably not a good explanation since none of the clusters chosen would qualify as non-transparent in the sense of the collocational continuum discussed above (see Schmitt et al. 2004: 130 for the list of 25 clusters tested). The eye tracker experiment, investigating whether the last word in a formulaic sequence received less attention than the rest of the sequence, pointed in the direction of the tested pieces of formulaic language being processed as wholes rather than analytically by native speakers. By the end of the sequence, they knew what to expect, so they did not need to fix their eyes on the final item for long. The results for non-native speakers were largely inconclusive; they reacted differently from the native speakers to the test as a whole, making comparability difficult. They generally displayed more eye fixations on the tested sequences, but the fixations on the final item in the sequences did not seem to be shorter. The final study in this series, using self-paced reading, showed the same results. Natives read faster than non-natives, but in terms of the sequences tested for whole-unit processing, no conclusions could be drawn. The authors speculate that the research methodology is invalid for answering whole-unit storage questions.

### 7.7 Perspectives – formulaic language, UBL, and SLA

In UBL, instances of formulaic language are symbolic units like all other linguistic units, so it is naturally questioned if these formulaic units present a specific problem for learners. The specificity of the learning problem, I would argue, has only emerged as a truism because it is seen against the backdrop of rule learning and therefore, when somebody happens to master this or that word combination but not other word combinations of the seemingly same kind s/he is thought to be linguistically moving beyond the capacity of her/his linguistic system. Of course a system cannot generate an item whose underlying rules it does not control, therefore the word combinations in question must reside outside the rule-system, so the argument goes. Keeping in mind the idea that we know words by the company they keep (Firth 1968, quoted in Poulsen 2005), I believe that my five research papers have shown the fruitfulness of describing SLA along principles laid out in UBL, to describe processes in SLA as item-based rather than system-based. Formulaic language, operationalised as recurring multi-word expressions used for a relatively coherent and stable purpose, was found in EC, ESK1 and ESK4, to be a good place to start in this respect because everybody, regardless of theoretical predisposition, agrees that it abounds in natural languages; it holds a natural position on Radden's continuum of symbolic units, as mentioned above, and it holds notorious relevance in more traditional approaches because it tends to defy general rules of language. Furthermore, formulaic language is frequently viewed as a problem-area in SLA because it does not readily fit into the system of grammar, nor does it fit the list of lexical items. Thus, questioning the status of compositionality as the norm, taking Formulaic Language seriously with all its implications, including the three points made above, and considering "the fact that we don't standardly combine our utterances all the way from minimal items to complete utterances" (Harder 2001: 234), I ask what *do* we do, then? What sort of 'unit' is language use based on, and by extension, what characterises the sort of 'unit' at play in L2 learning?

The question asked in the five research papers is if UBL's maximalistic model can exhaustively account for second language acquisition and use. It has been answered affirmatively so far. It does seem possible, in SLA, to describe learner language in terms of specific instances of language rather than an abstract system of rules. Taking into consideration the UBL idea that language learning is largely item-based rather than system-based, i.e. that people learn languages by abstracting away general properties from the exemplars they are exposed to, rather than acquiring an algebraic system of rules which is used to combine elements that carry meaning ('words'), I think it has shown

itself to be useful to explore the question of whether it is fruitful, in terms of SLA, to view language knowledge as a structured inventory of symbolic constructions rather than a list of words and a range of rules for combinatorial purposes. The most important truism to challenge, and which, seen retrospectively, scholars like Nattinger (1980) and Pawley and Syder (1983) seemed to be struggling with as well, is that language learning consists of learning semantically empty combinatorial rules and semantic units, *words*, with formulaic language constituting its own homogenous category in-between. UBL sees it fundamentally differently. All items of language are essentially made of the same fabric, a manifestation and a meaning. There is no compartmentalization of language knowledge. Tomasello (2003) summed it up (for children, but it may be transferred to be tested on adults); people learn two faces of grammar: smaller items and larger patterns. They may be actual instantiations of language, often-used routines and resources that we carry around to implement more or less at our will. They may be partially schematised, collocational frameworks (Sinclair and Renouf 1991), formulaic frames (Dabrowska 2000), utterance schemas (Tomasello 2000), consisting of fixed patterns and open slots, e.g. the *x'er the y'er*-construction, or pivot schemas (Tomasello 1992) with one fixed element controlling the rest of the pattern, e.g. *I can Verb*. They may even be fully schematised construction templates, e.g. the double-object construction *he baked her a cake*, or the caused-motion construction *he sighed the hand-out off the desk*, in which the construction template itself contributes to the meaning of the actual word-string (Taylor's coercion-principle, as discussed in the previous chapter). So the items to be learnt, according to UBL, are those that are afforded to us in interaction, starting with concrete items and evolving into an ever-changing increasingly schematized "construct-icon" (Goldberg 2003). Language knowledge in this conception is a moment-to-moment thing, something that happens while interaction and language use is actually taking place in usage events and something which, remembering Ellis' words, changes across the lifetime because it is, fundamentally, exemplar-based. "Language knowledge is a collection of memories of previously experienced utterances" (N. Ellis 2002: 166).

As pointed out in ESK1, some branches of recent work in corpus linguistics share some assumptions with usage-based linguistics such as the view of differences among language patterns as matters of schematicity and abstraction (e.g. Stefanowitsch and Gries 2003; Hoey 2007; Stubbs 2007). Especially worth mentioning here, however, because it has implications for the notion of 'collocation', is Hoey's lexical priming theory. Hoey (2007) holds that collocations form a

subconscious catalogue of frequent word combinations, primed by words in the combinations, and as such, he argues, they should be accounted for in psycholinguistic terms. In terms of acquisition, the implication is that the grammatical system emerges from the patterns encountered in use, the language patterns constituted by the collocations forming the basis of grammatical extractions in ontogenetic experience. This is parallel to the essence of the citation from N. Ellis above.

What I explore in my research papers is the application of such an experiential view of language learning to developmental issues in L2 learning with special reference to the role played by recurring multiword expressions. I do this by reference to UBL, a bundle of linguistic theories which aim not to dichotomise lexis and grammar, which see semantics as fundamentally non-compositional, and which do away with the rule/list fallacy, which results in a fundamentally different view of what constitutes the true unit of learning in SLA. Applying the UBL framework to the investigation of what I have operationalised as recurring multi-word expressions (MWEs) has empirically invalidated the viewpoint that such linguistic resources and routines are beyond the current interlanguage capabilities of the L2 learner. Rather, it would seem that given syntagmatic strings, constructions or utterance schemas depending on the degree of abstractness thought to underlie them, are item-based in acquisition, that MWEs may be locally recurrent and wander in and out of use, as it were; i.e., they need not be retained by the learner in order to qualify as a MWE. They come and go as environments change, they wax and wane in response to contextual requirements (Thelen and Bates 2003; Larsen-Freeman 2006). This has paved the way for an increasingly locally contextualized orientation towards linguistic knowledge; it has also resulted in the insight that what seems to constitute the bulk of portable linguistic resources for the individual learner is what I have operationalised as utterance schemas, i.e., partially abstract, partially concrete, lexically filled linguistic patterns. Together with MWEs, these seem to be the paramount linguistic resources and routines in operation for L2 learners.



## Chapter 8:

### **E pluribus unum? – tracing the roots of SLA**

#### 8.1 Introduction:

In recent years, the field of Second Language Acquisition (SLA) has evolved into a highly heterogeneous enterprise which embraces theories and approaches operating along the lines of a traditional cognitively oriented psycholinguistic vein, as well as more socially informed approaches, such as 'CA for SLA' and language socialization studies. Attempting to incorporate insights from both cognitively and socially informed research, I explore in the five enclosed research papers the notion of second language (L2) learning along a number of dimensions, which have only recently been brought to the field. Epistemologically, the point is not to reinvent and unify SLA as a field; rather, the dimensions along which I envision L2 learning studies to most profitably be undertaken will allow the field, not to be reconciled, but to explore a new eclecticism. These new dimensions are mostly a matter of operationalizing pre-existing dualisms as continua, most famously, probably, the performance-competence distinction, but also related ones such as the use-learning, and the sociality-individuality distinction. Discussions of these, and other dualisms, will materialize more clearly as this chapter traces the evolution of the field of SLA, the point being to situate the present research as captured in the five research papers in the general epistemology of SLA.

#### 8.2 There's competence and then there's competence – a brief Introduction to SLA's adolescence.

Tracing the roots of the field of SLA, two initial major influences dominate the picture: Behaviourism and generativism. Before Chomsky revolutionised linguistic science, among other things with his devastating critique of Skinner, most SLA theory and research was rooted in Bloomfieldian structuralism and Behaviourism. This early part of the field history is told in numerous places, e.g. R. Ellis (1985, 1994); Hakuta and Bialystok (1994); Ritchie and Bhatia (1996); Mitchell and Myles (1998/2004); Gass and Selinker (2001); Block (2003); and most recently VanPatten and Williams (2007); for the purposes of the present research, it need not be re-told in further detail here. Suffice it to say that early SLA research, as it was carried out by practicing teachers, was concerned with mapping out the structural differences between the mother tongue (L1) of the learners and the target language (L2) they were attempting to learn. This was referred to as *contrastive analysis*. The working hypothesis behind this endeavour was that

structural differences and similarities between languages would impinge on the learning process; where the languages differed the learners would encounter problems, as the habits of the L1 would *interfere* with the acquisition of L2 structure; where the languages were similar, the learners would not experience interference problems. Predictions could thus be made about learning trajectories, and errors, as demanded by behaviourist psychology, could be eradicated lest they would become bad habits. These undertakings were eventually abandoned because learners were not found to follow learning trajectories as suggested by the contrastive analyses. A further reason for this change was Chomsky, which will become clear as the chapter moves on to the focus of this section, namely what is commonly agreed upon as the identifiably homogeneous modern field of SLA (Larsen-Freeman 1991; Block 2003), which came of age with the introduction of such influential constructs as *error analysis* (Corder 1967; Ellis 1985) and *interlanguage* (Selinker 1972, 1992; Sharwood-Smith 1991).

A note of exegesis before proceeding: It is not the point of the present exposition to give an introduction to SLA as such. The point is, rather, to explore the roots of the identifiably modern field of research and to depict how the field has evolved since its early beginnings. The picture of SLA that will emerge is a rich one of theory proliferation in which researchers position themselves especially in terms of their orientation to issues pertaining to social and cognitive aspects of language and language learning<sup>80</sup>. By thus getting to the kernel of present-day SLA, I will be able to position the present approach as presented in the five research papers in relation to other strands of research in the field.

### 8.2.1 Interlanguage and psycholinguistics.

At the time of Corder's and Selinker's influential writings, Chomsky's devastating critique of the input- and repetition-oriented tendencies prevalent in behaviourist thinking and his corresponding

---

<sup>80</sup> This epistemological discussion also entails a meta-discussion of the benefits, or drawbacks, of theory proliferation. It lies beyond the scope of this chapter to go into details, but there is a tendency that those researchers who argue in favour of theory proliferation are inclined towards a socially oriented view of language and language learning (e.g. Block 1996; Lantolf 1996; Firth and Wagner 1997), whereas those who argue that SLA should be concerned with a limited set of phenomena, mostly psycholinguistic in nature, requiring a limited theoretical frame of reference, maintain a compartmentalized view of language and a cognitivist view of language learning (e.g., Gregg 2001; Long 2007). (See full see discussions in e.g., Beretta 1991; Long 1993, 2007; Block 1996; Lantolf 1996; Gregg 2001; Doughty and Long 2003; Jordan 2004). I would ultimately argue that theory proliferation is an empirical issue, and based on empirical evidence from diverse branches of SLA, incl. language socialization (e.g., Kanagy 1999), CA for SLA (e.g., Brouwer and Wagner 2004); emergentism (Larsen-Freeman 2006), and my own framework of applying UBL to SLA as presented in the five research papers, I would maintain that a multitude of theoretical frameworks is required to explain the varied, complex, and chaotic learning trajectories witnessed.

embrace of "a mental reality underlying actual behavior" (Chomsky 1965: 4) as the object of research still reverberated. L1 acquisition research at the time was characterized by the so-called morpheme studies, exploring the idea that there is a more or less predestined order of acquisition; that learners go through certain stages, or 'developmental sequences', when acquiring a language.<sup>81</sup> An idea to be dealt with in due course, it was to become dominant in SLA; we note at this point that modern SLA came into the world at a time when generative grammar with its rigid phrase structure rules was in its heyday and morpho-syntactic structures accordingly were given a very prominent position in terms of what constitutes language knowledge and hence what essentially and primarily constitutes the stuff of learning in SLA.

With the ideas of contrastive analysis rendered insufficient by the Chomskyan revolution, Corder's (1967) notion of *error analysis* (EA) and Selinker's (1972) notion of *interlanguage* (IL) were bred into a world of linguistics and language acquisition research that was geared towards psycholinguistics. Whereas contrastive analysis had dealt with target languages (i.e., complete systems of L1 and L2, respectively), EA and IL offered a framework for investigating language in development. The two constructs are interwoven and to a large extent inter-compatible. IL is a theoretical construct intended as a frame in which to account for language acquisition as a cognitively active process. IL theory assumes that learners' language is rule-governed and systematic at all stages of learning, and that this system is identical to neither the L1 nor the L2 system. When learners make errors it is a product of hypothesis-making and –testing, the system being in development. Errors are not a question of L1 – L2 differences (though sometimes they can be) – primarily, they are a question of the learner making hypotheses about the target language structure. Whereas the purpose of contrastive analysis had been to prevent errors, lest they manifest themselves as bad habits in the learner, EA and IL were thought of as constructs used to identify, classify, and explain errors (e.g., R. Ellis 1985). According to Corder (1967), such errors can either be systematic (competence-related) or arbitrary (performance-phenomena with no bearing on the underlying competence). Only the former, obviously, are part of the interlanguage in a given learner's mind.

---

<sup>81</sup> Brown's morpheme study is one of the most famous studies into *the order of acquisition* (R. Ellis 1985; Cook 1993). Some early studies, following The Identity Hypothesis, equalled developmental stages in the acquisition of L1 and L2, (see e.g., Larsen-Freeman and Long 1991; R. Ellis 1994).

Even though much in early modern SLA is thus ascribed to issues of predictability and mental equipment, a somewhat paradoxical trait of the field is the conviction that the learner is an active hypothesis-testing individual; how much creative construction can there be if acquisition is pre-determined? Going back to Corder and Selinker, it would seem that even though both, along with Chomsky and other psycholinguists, seek to investigate the mental system underlying behaviour (continuously the cardinal question for psycholinguistic SLA research, see e.g., Tarone 1983; R. Ellis 1994; Gass 1998; Doughty and Long 2003), they differ on the nature of underlying cognitive structures. With Corder's distinction between competence-based errors and performance-related mistakes, and the parallels he sets up between L1 and L2 acquisition processes, suggesting that both children and adults operate on 'transitional competences' in language acquisition, L1 or L2, he seems to take sides with Chomsky on matters of systematic stability and homogeneity (R. Ellis 1994). Selinker (1972), on the other hand, argued that the variability evidenced in individual interlanguages, especially in terms of fossilization, suggested that the cognitive mechanisms responsible for processes in L1 and L2 acquisition were not the same. The former, he argued, might be dependent on biologically endowed mental equipment designed to handle language, whereas the latter he argued to rely on general cognitive mechanisms. This dictum is important, indeed necessary, because of the variability found at intermediate stages in the developmental sequences (e.g., Tarone 1983; R. Ellis 1985; Larsen-Freeman 1991). This paradox of variability, seemingly never resolved in SLA, is embraced by Lund (1996) as an invitation to detach from the notion of pre-determined sequences for acquisition; when there is so much variation due to learners actively forming hypotheses about the target language, how can we possibly speak of fixedness of those sequences to such a large extent? This field anomaly was previously noted by R. Ellis (1985) who, although he did not call it an anomaly, saw the developmental sequences as an instantiation of a grand-sweep view of L2 acquisition which could level out minor individual differences. Selinker's 1972 paper had already hinted at interlanguage systems as being permeable and dynamic, suggesting a more open-ended nature than implied by Corder's (1967) 'built-in syllabus'. In later research (Selinker and Douglas 1985; see also Tarone 2000), a broadly construed notion of *context* was inferred to play a major role in developing interlanguages, even though the object of research still lay with cognitive processes and structures underlying behaviour. Also, even though Selinker (1972) noted that very few L2 learners attain native mastery of a L2, this continued to be used as comparative baseline for successful learning, also by Selinker himself who encouraged comparison

between the IL and TL systems; a tendency critiqued by Bley-Vroman (1983), Firth and Wagner (1997), and Kasper (1997).

This 'comparative baseline' issue, revolving around questions of systematicity, variability, and homogeneity, remains largely unresolved in SLA. From a Chomskyan perspective language knowledge is homogenous and perfect. This was adopted by Corder (1967) in his pre-interlanguage term 'transitional competence' which implies that the interlanguage system owned by the learner is always systematic in an homogeneous way (R. Ellis 1994). Selinker (1969; quoted in Bley-Vroman (1983: 1) was less clear on this, pointing out the possibly ambiguous nature of interlanguage systematicity. In 1990 there was a debate on this issue with Gregg (1990) and Long (1990) arguing in favour of Chomskyan systematicity and Tarone (1990) and R. Ellis (1990) proposing a model of variable competence for L2 learning. This latter model, partially inspired by the insight from sociolinguistics that languages are not homogeneous, deliberately expanded the notion of competence to simply accommodate variability, instead of shunting it to the mess of performance. This allows for both inter-learner variation and intra-learner (i.e., interlanguage) variability. As such the approach suggested by R. Ellis and Tarone is compatible with my research as presented in the five research papers to the extent that both strands suggest that learning is a reflection of language use (R. Ellis 1994). Also, the issue of competition among rules in interlanguage development, as suggested by R. Ellis's variable competence model, is similar to the competition of structures as evidenced in my data (EC; ESK4).

Today, there is an incipient recognition in psycholinguistic SLA that individual differences, in terms of both end-points and intermediate stages in development, are considered too great to be ignored (e.g., Larsen-Freeman 2006; de Bot et al. 2007), and the notion of context, revitalized in Firth and Wagner (1997) in terms of social interaction, is now, following the calls from e.g., Kasper (2004) and Wagner (2004), also being investigated longitudinally to explain L2 learning (e.g., Hellermann 2006, 2007; Firth and Wagner 2007). As referenced in my research papers, also psycholinguistic researchers, proposing an emergentist outlook on L2 learning as a constant and iterative process, emphasize notions such as developmental non-linearity and multi-modality of language and language learning, suggesting that SLA explanations may be found in terms of social interaction, and not exclusively in terms of cognition (e.g. N. Ellis and Larsen-Freeman 2006; Larsen-Freeman 2007). The contribution from the present research, that which fundamentally sets it

apart from the variability model referenced above, is that the nature of linguistic 'rules', as also argued in ESK4, is viewed differently in R. Ellis's model and the present framework; whereas Ellis speaks of generally applicable rules of syntax used cross-constructionally, I specifically speak of construction-specific patterns which are mini-grammars (Pawley and Syder 1983) onto themselves. The emergent system, as proposed in Usage-Based Linguistics, is then the coming together of many such mini-grammars, which eventually allows for the abstraction of regularities and schematised constructions in the internalised inventory of linguistic knowledge.

Such new ideas will be discussed in detail later; for now, it may be inferred that the coming of age of the field of SLA is steeped in psycholinguistics<sup>82</sup>. Furthermore, epistemologically, the field of SLA as initially and, in fact, traditionally conceived, shares a number of fundamental viewpoints and research objectives with Chomskyan linguistics (Firth and Wagner 1997). Among these are the presence of structural universals of language, a strong mentalist and individualist orientation, and the notion of fixed developmental sequences in acquisition. Even though environment has a part to play, too, in that input is seen as a decisive factor in L2 acquisition, SLA theories that do not account for formal linguistic universals, the contribution to processes in L2 acquisition of specialized language acquisition equipment in the mind of the learner, and adherence to regular and systematic interlanguage rules and sequences are deemed incomplete or inadequate at this stage in the development of the field (Long 1990). Even though Lund (1996), in an attempt to apply Givón's (e.g., 1985, 1995) functional linguistics to SLA, found empirical evidence against the fixedness and the universality of developmental sequences, they remain to this day (Pienemann 2003; Unsworth et al. 2006; VanPatten and Williams 2007) a part of the established findings that helped shape SLA as a field of enquiry onto itself.

### 8.2.2 Developmental sequences

A variation, perhaps a constriction, of the idea of the active hypothesis-testing individual language learner is Pienemann's (e.g., 2002, 2003) Processability Theory (PT) according to which the incremental manner in which interlanguage progresses along the pre-paved road of the

---

<sup>82</sup> Even though Chomsky played an important role in focussing linguistic inquiry exclusively on cognitive/psychological issues, not all psycholinguistic SLA researchers are explicitly Chomskyan (Long 1997) (but see e.g., Cook 1985; Gregg 1990; Carroll 2002, 2007; White 2007, for researchers working in a Chomskyan vein). However, the strength of Chomskyan influences can be seen in Krashen's Monitor Theory, arguably the first full-fledged theory of SLA (Block 2003; VanPatten and Williams 2007), which, fundamentally applying Chomskyan thinking to SLA, in essence counts as an Identity Hypothesis equalling L1 and L2 acquisition processes and positing fixed developmental sequences for the acquisition of syntactic rules.

developmental sequences must be restricted not only by aspects of linguistic complexity, but also by constraints that are psycholinguistic in nature (similar in nature to Corder's inbuilt syllabus, as already mentioned). Certain structural aspects of language, according to PT, may only be learnt and thus profitably taught, at certain points in time, depending on the learner's present progression on the implicational scale of structural language acquisition. This puts natural limits, or constraints, on the options of hypothesis-testing; structures too complex to handle for the present level of interlanguage competence cannot be processed by the learner.

Learning, according to PT, essentially starts from unstructured bits and pieces. Learning proceeds along a line of increasing structural complexity: words → category ascribing (grammatical) → intraphrasal manipulation → interphrasal manipulation → subordinate clause structure manipulation. The objects of learning, then, are all structural. Category ascribing concerns pluralizing nouns, for example, or tense-conjugating verbs. It is inferred that manipulation of the items presupposes category ascribing. The intra- to interphrasal manipulation has to do with the distance between structurally related linguistic information; learners acquire structural relations which are relevant inside phrases before those that are relevant across phrases<sup>83</sup>. The relationship between *two* and *men* in the NP *two men* is acquired before the relationship between *he* and *comes* in the sentence *he comes*, which requires the production of an agreement relationship between a NP and VP. Interphrasal manipulation ability is followed by the ability to master subordinate clauses, to the extent that these differ structurally in the TL (see Glahn et al. 2001).

In terms of the exemplified trajectories outlined below, stage  $x+1$  includes the ability to move a limited number of items around but without being able to manipulate the structures themselves; i.e., if an adverb is put before the main structure, the default SVO-word order – the learners' original hypothesis, it is assumed – is retained even though the target language rules require inversion. At the next stage, intra-phrasal manipulation becomes possible (which is necessary to master the verb separation structure in German and the pseudo inversion for Y/N-questions in English, see below). The grammatical exchange between phrases, as described above, is seen to coincide with Inversion manipulation abilities, and ultimately the final level is reached where the learner masters

---

<sup>83</sup> The notion of *phrases* refers to phrases in a Chomskyan vein. These are centered on the major syntactic categories Verb, Noun, Adjective, and Preposition, which are thought to be candidates for universal features of language. They are usually referred to as lexical phrases, because their heads, e.g., the Noun in a Noun Phrase, are drawn from the lexicon. They have nothing to do with Nattinger and DeCarrico's (1992) lexical phrases as multi-word items.

subordination and other complex syntax operations. The acquisitional hierarchy thus outlined is seen as implicational; the linguistic constraints in play in language acquisition are evident in the gradually more complex syntactic operations underlying each stage on the developmental course, and the psycholinguistic constraints mirror these; it is impossible to process any linguistic information on a level beyond the level of the current interlanguage system. As Pienemann (2002: 190) put it, "[processability theory] accounts for the trajectory of L2 development on the basis of the constraints that derive from the architecture of the emerging L2 grammar." No sequence can be short-circuited, none of the levels skipped.

Dealing exclusively with morpho-syntactic aspects of language, early research, the so-called ZISA-project<sup>84</sup>, on developmental stages for general word order acquisition (Pienemann 2003; Pienemann et al. 1988) and inversion acquisition specifically (Pienemann 1985) in German SLA, pinpointed the following stages in development:

- Stage x      Canonical order (SVO being the learners' initial hypothesis)
- Stage x+1    Adverb Preposing (without obligatory inversion, German being V2; i.e., the learners stick to the canonical order of their initial hypothesis)
- Stage x+2    Verb Separation (obligatory in German; SVOV)
- Stage x+3    Inversion (after "preposing of elements", so Pienemann et al. 1988: 221)
- Stage x+4    Verb Final (finite verbs move into final position in subordinate clauses)

In terms of application of the inversion, learners on Stage x+1 have not come to mastery, neither in declaratives with adverb preposing nor in interrogatives. On Stage x+2 the same picture emerges; however, the learners may have inversion in Y/N-questions. On Stage x+3 inversion is in place in all obligatory contexts, and x+4 involves the ability to handle the most complex structural phenomena, namely differences between main and subordinate clauses, which require inter-phrasal manipulation and awareness of the structural differences just mentioned.

According to Pienemann et al. (1988: 222), the developmental sequences constitute "one of the most robust findings in SLA research". The vast number of studies examining SLA in the light of those sequences seems to speak in favour of their viewpoint. The developmental sequences and PT

---

<sup>84</sup> Zweitspracherwerb Italienischer und Spanischer Arbeiter



studies and research overview papers include Hyltenstam (1977, 1985); Meisel et al. (1981); Pienemann (1981, 1985, 2002, 2003, 2007); Clahsen (1985); Clahsen and Muysken (1986); Pienemann et al. (1988); Håkansson and Nettelblatt (1993); Pienemann and Håkansson (1999); Glahn et al., 2001; Håkansson 2005; Unsworth et al. 2006). Furthermore, quite influential, these studies are featured prominently in many introductory textbooks in the SLA field; e.g., Ellis (1985, 1994); Larsen-Freeman and Long (1991); Lightbown and Spada (1993); Mitchell and Myles (1998/2004), and very recently VanPatten and Williams (2007).

Mackey (1999), in an investigation of question formation development also used Pienemann's developmental stages to determine progression. In other words, she did not test the validity of the sequences, she used them as comparison baseline. Parallel to the sequences established for German SLA, as outlined above, the path for English L2 question formation acquisition includes the following developmental stages<sup>85</sup>:

- Stage x      Canonical order (SVO being the learners' initial hypothesis; Mackey adds that the learners may apply question intonation. Example: *it's a monster?*)
- Stage x+1    Fronting *Wh/Do/Q*-word (example: *where the cats are?, do you have an animal?*)
- Stage x+2    Pseudo inversion: Y/N, Copula (example: *have you got a dog?, where is the cat in your picture?*)
- Stage x+3    Do/Aux-second (example: *what do you have?*)
- Stage x+4    cancel inversion in subordinate clauses (examples *can you tell me where the cat is?*)  
Neg Q, Tag Q (examples: *haven't you seen a dog?, it's on the wall, isn't it?*)

Studies in developmental sequences deal exclusively with morpho-syntactic aspects of language (Clahsen 1985; Hyltenstam and Pienemann 1985). Even if there is no a priori theoretical consensus among all researchers working within this tradition, it would seem that it finds Chomskyan kinship in at least two respects. One is the shared syntactocentric focus, that syntax may be studied independently of other components of language, implying autonomy and modularity, the other is the idea that external influences play no role in acquisition, rendering the developmental sequences universally pre-determined. For instance, Hyltenstam (1977: 408), in his examination of the

---

<sup>85</sup> Mackey lists the stages numerically as stage 2 through 6. In Pienemann et al. (1988), one of the original papers to posit this developmental route, stage 1 consists of formulas. Mackey, apparently, does away with this stage. I have retained the listing format from above in the interest of comparability.

acquisition of Swedish negation syntax, i.e., negation placement rule, found learners to travel identical routes of acquisition, which, he argued, "points to a confirmation of a view of universally determined language development."

This indisputable, universal relevance seems to be well-established among these researchers; Hyltenstam (1985: 120) states that their reality is "a well-established fact", Pienemann et al. (1988) called them a robust finding, as mentioned, and according to Meisel et al. (1981: 110), it is "assumed by many authors that [developmental stages are] normally the same for each individual learning a second language". However, Meisel et al. (1981: 110) also claim that "one has to distinguish between different groups of learners who may follow different paths on their way to the target language", thus allowing for variation. Looking a bit more closely it transpires that other researchers (e.g., Clahsen 1985; Pienemann et al. 1988) allow for variation too. R. Ellis (1984), discussing the acquisition (order) of the pronouns *who*, *what*, *where*, and *when*, goes even further in terms of variation, pointing out that formal instruction can change the route or speed up the process, of acquisition, which is in direct opposition to Pienemann's PT.

Not entirely convinced of the validity of the universalism of the developmental sequences, R. Ellis, as already noted, later went on speculate whether they were, more than anything, just an overview of morpho-syntactic developmental tendencies, bleaching out the colours of variance. That early scepticism can be seen as preliminary to R. Ellis's own variable competence model (as also propagated by Tarone (1990) and as such he became a forerunner of later scepticism as voiced along the same lines in N. Ellis and Larsen-Freeman (2006) and de Bot et al. (2007). In ESK4, I argue that the developmental sequences do not fit the data from my focal students, and I further argue that the underlying language views of UBL and the formalist tradition behind the developmental sequences differ to such an extent that the acquisitional insights drawn from research are different, irrespective of the status of interlanguage developmental sequences.

The issues marking the differences between my research and PT research, pertain to the establishment of a point of acquisition of the phenomenon under investigation. Working from the perspective of a 'construction grammar' (e.g., Fillmore et al. 1988; Goldberg 1995; Croft 2001; Fried and Östman 2005; also outlined in the previous chapter) I investigate the acquisition of syntagmatic relations on the basis of recurring patterns. My focus, in other words, is on patterning

of language and the extent to which my data indicate that my focal students have learned separate patterns (e.g., *I don't know*) or more schematised linguistic knowledge (e.g., *I don't Verb*). As also pointed out in ESK4, I do not work on the assumption that structural features are acquired on a pattern-independent basis; rather, I assume – and my data have empirically substantiated this assumption – that structural features are learned on a pattern-by-pattern basis. The learning of these features, then, are dependent on pattern acquaintance, including aspects of frequency. This was illustrated in ESK4 by Valerio's learning of the *you no Verb*-pattern and the subsequent competition against the target-like variety, *you don't Verb*.

Researchers working within PT, on the other hand, because they focus on cross-constructural structural features, usually invent arbitrary frequency thresholds for acquisition, i.e., in order to count as acquired, a given structural feature must be applied in a certain number of contexts relative to the total number of obligatory contexts found in the data. For example, Glahn et al. (2001), who investigated (and reported to find general support for<sup>86</sup>) Pienemann's *processability theory*, set this number at 50% or 80% and the results are then compared. In terms of operationalizability this might be sensible, but to determine whether or not a generically applicable form is 'acquired' as part of the linguistic system based on a 50% or 80% obligatory occurrence is completely arbitrary. They also work on a 'once produced, acquired' criterion, but basically abandon it because they see it as a problematic indicator of an acquired system, which, of course, is wise if you buy into the underlying premise. Arguably, however, it is only problematic because language knowledge is seen a priori as systemic, stable, and omnipresent rather than probabilistic, emergent, and locally situated; it could also be said to be a consequence of the belief that language learning is based on structural implicational scales as presupposed by PT.

Furthermore, for the purposes of the present research, applying the %-methods, whether 50 or 80, is problematic, because it requires an elimination of MWEs. This MWE elimination is necessary because MWEs usually display correct production of the morphological or syntactic phenomenon under investigation. Consequently, MWEs are often, as mentioned in the previous chapter on

---

<sup>86</sup> It is beyond the scope of this chapter to undertake this discussion in detail, but Glahn et al (2001) investigate cross-sectionally the acquisition order of three formal relations in learners of Scandinavian L2. Their cross-sectional data do support PT, but if the data are scrutinized informant by informant, it becomes clear that individually assessed, people do not necessarily abide by Pienemann's processing laws. Some of the informants do, but the fact that some do not is sufficient for me to find the theory insufficient at best, invalid at worst. In any case, my personal preference is not to have the broad view of structural acquisition, but to investigate the nitty-gritty detail of individual ontogenesis in a L2.

formulaic language, considered to be beyond the current interlanguage capability of the learner. If there is a high recurrence of a MWE, it might yield skewed results suggesting the system to be in place. Obviously, given the interest in the present research in such MWEs, this *modus operandi* is useless. Two things may be noted, then: 1) the underlying views of the 'system', in PT and the present approach, respectively, are incompatible; and 2), as also noted by Bardovi-Harlig (2002), it shows how atypical it is in research in developmental issues in SLA to take MWEs as one of the points of departure.

While my research thus fundamentally differs from that proposed in PT and developmental sequences research, other cognitively oriented researchers have criticised the model for not explaining in more exact terms how processability limits acquisition (Larsen-Freeman and Long 1991). More in alignment with my research, Skehan (1998) criticised PT for largely ignoring lexis and for shunting the notion of 'chunks' to the periphery. Skehan (1998: 39), proposes that learners "have to become more lexical in their mode of communication, and correspondingly in the repertoire of language knowledge that they possess", which is regarded with sympathy here. However, his view of a cognitive information-processing model builds on a compartmentalized view of language knowledge which is fundamentally incongruent with the views espoused in UBL. Skehan (1998) thus speaks of a dual-mode system of processing which in on-going communication rests on memorised chunks of the sort implied by the Idiom Principle, as already discussed in chapter 7 as well as EC and ESK1. Only in cases with few stress factors – time, pressure etc. – does the system switch to a syntactic mode of processing.

All SLA explanation, as explored so far in this chapter, is found in terms of cognition – and the cognitive representation of language knowledge is typically captured in a compartmentalized fashion along the lines of Skehan's dual-processing model<sup>87</sup>, cf. also Wray's (2002) influential working definition of formulaic language, as discussed in chapter 7. 'Environmentalist' views of SLA (Long 1990; Larsen-Freeman 1991) remain obscure; the challenger's viewpoint (Larsen-Freeman 2007). In hindsight, it is thus possible to detect the seed of the division of the field into a

---

<sup>87</sup> A notable exception is MacWhinney's (e.g., 2001) competition model, which subscribes to a connectionist view of language learning and a view of language knowledge which does not compartmentalize linguistic knowledge. The competition model, while not very apt as theoretical framework for dealing with interactional classroom data like mine, provides a framework for delineating the relative importance of certain phenomena ('cues') in the linguistic input. These cues are e.g., syntactic, morphological, semantic, and pragmatic ones competing for the learner's attention. In the sentence *the spoon kicked the horse*, figuring out who did what to whom can be done on the basis of word order (spoon is agent), animacy (the horse is more likely to be the agent), or case (impossible).

sociolinguistic and a psycholinguistic branch (Markee and Kasper 2004; Firth and Wagner 2007); the former would eventually find it very difficult to reach through to the other side with the message that perhaps the field, due to its adolescence, had been somewhat biased, tuned to individualistic, cognitivist, syntactocentric viewpoints concerned purely with linguistic competence. These are all matters to be further discussed below. For now, the following section pays a visit to yet another kind of competence, namely communicative competence. While there are many sympathetic thoughts for the socially oriented scholar in Hymes's (e.g. 1967, 1972) original writings on an ethnography of speaking and communicative competence, the field of SLA managed to mould an emergent change in the world of linguistics in its own image. Communicative competence was to fall victim to psycholinguistic tendencies to compartmentalize linguistic knowledge and distinguish it sharply from use.

### 8.2.3 Communicative Competence

Lack of theoretical consistency within the field of Communicative Competence (CC) has been noted by numerous scholars, e.g., Canale and Swain (1980); Canale (1983a, 1983b); Færch et al. (1984); Schachter (1990); Celce-Murcia et al. (1995); Brown (1996); Riley (1996); Shohamy (1996). This has to do with terminological distinctions concerning notions such as competence, performance, knowledge, skill, and ability for use, as well as uncertainty of the validity of the components of the CC framework. All CC frameworks do not share exactly the same sub-components, and the sub-components that they do share are not always delineated identically (Larsen-Freeman and Long 1991). The frameworks outlined in this chapter all seem to share three superimposed notions, as it were, namely, *competence – ability for use – performance*. It is not, however, entirely clear if *ability for use* permeates all sub-components of the frameworks or if it only pertains to areas of sociolinguistic, pragmatic and strategic competence, rendering linguistic/grammatical competence a knowledge-only component. These are all matters to be explored in the following.

Relatively early frameworks such as Canale and Swain (1980), Canale (1983a), and Færch et al. (1984) consisted to varying degrees of linguistic, pragmatic/sociolinguistic/discourse, and strategic competence. (Discourse competence was added as an independent sub-component by Canale; in the former model it belonged to sociolinguistic competence). *Grammatical* competence has to do with lexical items, rules of morphology, sentence-grammar semantics, and phonology. Sociolinguistic

competence is about sociocultural rules of use (appropriateness, context, participants, setting) and rules of discourse (cohesion and coherence). Strategic competence contains the tools compensating for breakdowns, typically lexical in nature (see also Paribakht 1985). It will be noted that the lexicon is given a minor role in grammatical competence; arguably, in a compartmentalized view of language it would deserve more attention – its own compartment. In Færch et al.'s model, linguistic competence is embedded in pragmatic competence which in turn includes the sociolinguistic component, with strategic competence as something to be employed in instances of communicative failure.

These scholars do not share precise categorizing principles in terms of the sub-components, but the basic idea of linguistic rules of language and pragmatic rules of use, with strategic competence as a tool to be employed in compensating for breakdowns, mostly lexical in nature, remains largely identical.<sup>88</sup> Epitomised by Schmidt (1983: 172) who discusses the "partial independence of Grammatical Competence from other components of Communicative Competence", the idea of dividing language and knowledge of language into sub-components in the vein of CC reads like a variation on the theme of compartmentalization. In other words, it seems that the focal point is underlying Competence, not actual use; as pointed out by Skehan (1998) if we employ CC as a means to measure Performance it is nothing more than a 'checklist'. Considering the fact that CC is an alleged attempt to fundamentally change views on language knowledge, this similarity, to say the least, is not favourable. Yet it does not end here. The relatively sharp distinction between Competence and Performance, with the inter-mediating *ability for use*, resembles the Chomskyan model of *Grammatical Competence – Pragmatic Competence – Performance* prevalent at the time (Chomsky 1980) of the initial application of CC to SLA.

Building on his dictum that rules of language are useless without rules of use, Hymes (1972) originally wanted to stress the complexity, multi-modality, and variance of language use and speech communities. While this may at first seem to be another add-on to the construct of linguistic competence, it is, in fact, a radical riot against it. 'Rules of use' was a useless notion to Chomsky. The CC formulation by Hymes rests on an insight that the Chomskyan notions of competence and

---

<sup>88</sup> This conception of strategic competence is in alignment with scholars dealing with communication strategies, e.g., Paribakht (1985); Poulisse (1993); Kasper and Kellerman (1997); and Singleton (1999) who all attend to lexical difficulties and see communication strategies as primarily a remedy for referential problems focusing on lexical knowledge. It seems paradoxical, then, that lexis as such, included as a minor element in linguistic competence in the early frameworks, is afforded so little attention, a problem mentioned above.

performance did not adequately account for language use, sociolinguistic insights should have a constitutive part to play in linguistic theory (Hymes 1972). Perhaps due to Hymesian influences the Chomskyan tradition has become fully aware of the need for a theory of language use to complement their theory of Competence (Chomsky 1980; Cook and Newson 1998; Taylor 1988). Hymes's ethnography of speaking, however, came into the world at a time when other approaches to the study of language in use, or language as rooted in everyday behaviour, were emerging such as ethnomethodology (Garfinkel 1967) and Conversation Analysis (Sacks et al. 1974). Arguably, the latter became the most prominent research movement applied in studies of the systematicity of real-time interaction with its discovery that conversation is governed by a systematic turn-taking machinery (Sacks et al. 1974). Such approaches view language, and language and learning, in an entirely different light, as will be discussed below.

Hymes's sociolinguistic approach to issues that had to do with the complex skill of using language was, in SLA's terms, superimposed on compartmentalized view of language knowledge. Even though Hymes (1967, 1972), as a student of social life and social interaction, argued that the place of linguistic theory was in a sociocultural theory and specifically claimed that language structure and language use are interdependent (Hymes 1972), the CC framework, for SLA purposes, came to be understood as "underlying systems of knowledge and skills required for communication" (Canale, 1983a). Even though this was a reaction to Chomsky's limited notion of linguistic competence – as Færch et al. (1984) note, CC marked a shift of emphasis from the traditional Chomskyan focal point of internalised rules towards *language in use* – the failure to ultimately detach from a Chomskyan approach to language is summarized in Canale and Swain (1980: 6) where it is held that CC is *knowledge* (of rules of language and use) and performance is realisation of CC. This only differs in nature from the Chomskyan framework of linguistic competence, pragmatic competence, and actual use to the extent that it seems to conflate linguistic and pragmatic competence. The performance-competence distinction is intact.

Thus, when other scholars embarked on research in language use, problems have arisen as they have applied Chomskyan terminology to their models. Hymes has also been misunderstood (see discussions in e.g., Taylor 1988; Widdowson 1989; McNamara 1995); his version of competence was a deliberately expanded notion which included knowledge, ability for use, and instances of use. Expanding the 'knowledge notion' Hymes, at face value, has been a pioneer, both in terms of the

present research as put forward in the five research papers, as well as in recent writings of e.g. Larsen-Freeman (2007) and other, more socially oriented SLA research (Firth and Wagner 1997). Hymes (1972) also opposed himself to Chomskyan homogeneity principles, arguing that it cannot automatically be assumed that all speakers of a language have the same linguistic knowledge. It is in this respect that the variability model (R. Ellis 1990; Tarone 1990) have inherited sociolinguistic insights. Sadly, such fundamental insights seem to have been drowned in the attempts to have the CC model comply with a pedagogical checklist model (Skehan 1998).

The major point to be made in this section is that Hymes's contribution to the ongoing debate was important in that it took emphasis away from rules of language to rules of use. For acquisitional research purposes, however, it did not change the conception of the nature of the underlying competence to be acquired. Hymes (1989) is very clear on the use-knowledge issue himself, arguing that to be competent in a language requires an ability, not pure knowledge (i.e., mental representations) alone. In fact, use and knowledge virtually coalesce in Hymes's original writings (Gumperz and Hymes 1972). Furthermore, not only does he differ from Chomsky in terms of infusing language use with theoretical and empirical importance, he later went on to state explicitly that his views of language knowledge and acquisition parallel those of functional linguistics and connectionism (Hymes 1989). Thus, even though as noted, 'CC for SLA' did imply a hitherto unseen focus on actual language use, the underlying compartmentalized apparatus for researching SLA remained unchallenged and unchanged. To quote Hymes (1989: 245) himself, it was as if "new horizons were meant to be seen from certain windows" only. "Hymes retained the idea of underlying competence", so Skehan (1998: 157) even though it seems quite clear from Hymes's original writings that he had something entirely different in mind (see also Widdowson 1989). Skehan even goes so far as to denounce CC frameworks in SLA theory for the simple reason that there is no LAD at play in second language learning. He might have a point but that is not Hymes's fault; responsible are those who initially applied communicative competence to SLA.

*Quod erat demonstrandum:* The language view reflected in SLA's adoption of CC theory is a far cry from the ideas proposed by Hymes himself. Even though Widdowson (1989) points out that Hymes's and Chomsky's views on language are not commensurate, Hymes being interested in language, Chomsky in grammar, the field of SLA coerced the meaning of CC to fit its pre-existing agenda of investigating the incremental acquisition of morpho-syntactic rules along the laid-out



interlanguage trajectory. SLA's roots in psycholinguistics with its compartmentalized view of language knowledge, a total separation of competence and performance, and thus of acquisition and use, proved too strong to succumb to new, more socially oriented ideas; ideas which would potentially open up the field towards an embrace of the real world of language users; as Larsen-Freeman (2004) notes, social views of language are now moving towards fuller participation in the community of SLA researchers. At the time, however, the focus remained firmly, and solely, inside people's heads; there was competence and then there was competence.<sup>89</sup>

### 8.3 Individual competence or social performance?

As might be inferred, then, a retrospective glance at what happened in the field in the 80s parallels to a great extent much of what is being debated in present-day SLA (Firth and Wagner 1997, 1998, 2007; Block 2003). Firth and Wagner (1998) invited the field to open up spaces (Watson-Gegeo 2004) to include a more fine-grained view of interaction – only to be dismissed as sociolinguists who were not really interested in acquisition, only use<sup>90</sup>. And much like Hymes, they argued that the two should be thought of as inseparable. But not so to the field at large; use goes on out there, acquisition takes place in here, and never the twain shall meet. I am getting ahead of myself here though; I will return to the discussion of a field divided into sociolinguists and psycholinguists (Markee and Kasper 2004). For now, it will be noted that linguistic development in traditional SLA has been thought of as detached from interactional biographies of learners (Ohta 2001). This tendency has been pointed out by Tarone (2000), who noted that social contextual aspects of L2 development are largely uncharted territory, as well as Kasper (2004: 551), according to whom we do not know much about "the relationship between interlanguage knowledge and the environments in which such knowledge is first constituted, developed, and made increasingly more available for effective L2 use." That is a tendency which I aim to change in ESK1, ESK2, ESK3, and ESK4, and I think that the UBL framework with its assumption that language learning is a matter of abstracting constructional regularities in usage events is particularly apt at capturing precisely the interplay Kasper (2004) talks about.

---

<sup>89</sup> Interactional competence (Kramsch 1986; Hall 1993) is the concern of a later section due to its affiliation with socially anchored views of language and language learning.

<sup>90</sup> Long (2007), in an aggravated denial of the relevance of any ideas put forward by the likes of Firth and Wagner, Lantolf, and Block, repeated the dismissal of what Larsen-Freeman (2007) referred to as the challenger's viewpoint which has shown itself worthy of consideration. Long (2007) even goes so far as to question the soundness of the editors who published papers by Firth and Wagner (e.g., 1997) and Lantolf (1996).

At this stage, echoing Larsen-Freeman (1991); Block (2003); Gass and Mackey (2007), we are able to define traditional SLA as a field of research interested in learning and the learner's underlying linguistic system and the cognitive processes involved in the organization of this system. Apart from Skehan's suggestion that learners' need to be more aware of lexical chunks of language, however, the discussion so far has not had much to say about the nature of that which the learner learner's cognitive machinery operates upon. Hence, this section is concerned with the notion of *input*.

### 8.3.1. Is there a super-theory of SLA?

It is maintained in psycholinguistics that the existence of language items in long term memory (LTM) is a prerequisite of fluency (McLaughlin et al. 1983). Following Bates and MacWhinney (1982), avoiding to pose linguistic constructs that have no psychological validity, the question which the cognitive approach (Skehan 1998) must answer is what these language items are, what categories linguistic knowledge consists of. It seems an established fact in cognitivist SLA that the mental representations of language, whatever their nature, comes from the *input*, which is *attended* to and *noticed*<sup>91</sup> by the learner (e.g., Sharwood Smith 1986; Gass 1988; Schmidt 1990, 2001;) and given enough attention, language items may become stored in LTM. This implies a difference between input and intake (Corder 1967; VanPatten 1990), the latter defined as a subset of the former, that which the learner actually comprehends and operates upon. Incomprehensible language input cannot be 'taken in' (Long 1983; VanPatten 2007). This in essence is the shared viewpoints in various psycholinguistic interaction hypotheses, as they have been proposed in various guises in psycholinguistic SLA by "authors ranging from Peter Skehan to Rod Ellis to Susan Gass" (Block 2003: 92); items encountered in the *input* are subject to *attention* or *noticing* by the learner's cognitive processing system in order to become incorporated into the competence system which is then *restructured* (McLaughlin 1990).

There are also aspects of quality of input involved; saliency and frequency aspects of input influence intake (N. Ellis 2007), but often there is little discussion of the very nature of the input and hence the language knowledge to be attained. Schmidt (2001: 30-32) tackles this issue, listing phonology, vocabulary, pragmatics, morphology, and syntax as areas to be attended but there is not much in terms of details. The closest we come is the conception that "going beyond purely

---

<sup>91</sup> Noticing, a term coined by Schmidt (1990), also implies 'noticing the gap'; i.e., the gap between the current Interlanguage system and the target language system.

formulaic use requires that utterances be syntactically analysed". As will be remembered from Chapter 7, this is parallel to Wray's working definition of formulaic language which was argued to reflect a compartmentalized view of language knowledge; a separation of psycholinguistic categories into 'syntax', 'morphology', 'words' etc. In this respect, much psycholinguistic SLA research is in compliance with latter-day Chomskyan linguistics, perhaps most aptly captured in Pinker's 1999 book title *words and rules*. Recently, it will be noted, researchers such as de Bot (de Bot et al. 2007), N. Ellis, and Larsen-Freeman (N. Ellis 2002, 2003, 2004, 2007; N. Ellis and Larsen-Freeman 2006; Larsen-Freeman 2006, 2007), have been very clear on the issue of language knowledge, arguing in favour of a dynamic systems approach which is more apt at capturing the waxing and waning of linguistic patterns (Thelen and Bates 2003) as encountered in L2 data. Furthermore, N. Ellis (2002) explicitly called for the application of UBL to SLA, and as such he is, alongside Tomasello (2003), the most influential researcher in terms of theoretical footing of the present research. This is clear in all the five research papers, but most prominently in EC and ESK1.

Getting back to the issue of comprehensible linguistic input; the source of that input is defined with recourse to interaction between native speakers (NSs) and non-native speakers (NNSs) as modelled in the Input-Interaction-Output model (IIO; Block 2003). Early forerunners of the IIO model were Hatch (1978), who called for more focus on conversational interaction among L2 learners, and Krashen (1978), who argued that *input* was not only a necessary condition but in fact a sufficient condition for L2 learning. This was later modified by Swain's influential work on the importance of language production, i.e., *output*, by L2 learners (e.g., Swain 1985). The beginnings of the IIO-model as it stands today can be traced to the early 80s and the work by M. Long (e.g., 1983, 1985) who summarized, in the Interaction Hypothesis, prevailing viewpoints on and assumptions about conversational interaction, input, and output and their respective roles in SLA (Mackey 2007). The basic tenets of the model are that language goes in, is handled by some form of mental equipment as meanings are necessarily negotiated because there is a NNS involved, is tested against hypotheses about how the language in question is structured, and (new) language comes out; "new" is parenthesised because sometimes output is identical to or copied from input as echo (Gass 2003). Apart from that and in tune with the overall mentalist orientation, whatever takes place in interaction in the form of identity maintenance, power relations, or, as is done on a daily basis of people's everyday lives, "the organization of social action" (Schegloff 1988: 136), is not seen as central to processes of acquisition. According to the IIO-model, it may be inferred, people do not

really interact; they transmit knowledge representation in the shape of linguistic forms, one individual brain to another. The environment may be necessary as context, but sources of acquisitional explanation are located firmly inside the mind of the individual learner based on attention to and comprehension of the input. Attention to and processing of input may, if comprehended, lead to intake (Schmidt 2001; VanPatten 2007), which manifests itself in the interlanguage which is then restructured (McLaughlin 1990). Intake, or uptake, is the subset of input which defines what is learnt; i.e., taken in / taken up by the learner, primarily through negotiation for meaning in interaction.

In research papers ESK1, ESK3, and ESK4 I explore in some detail the possibility of finding interactionally oriented explanations for L2 development. I do this, however, not by focusing on the learner as being somehow attentive to certain linguistic items in the input, but by looking at recurring practices in the classroom. I do not see the distinction between interaction and learning very clearly in my data, and hence I cannot keep the two separated by inserting the notion of 'uptake' between them. This is the *acquisitional* tenet behind my impetus not to work on the IIO-model here; the *linguistic* impetus is equally important to the extent that IIO is generally affiliated with a formalist (UG) oriented approach to language structure (Block 2003) – or at least one which maintains a compartmentalized view of language knowledge (Gass and Selinker 2001) with a heavy inclination toward investigating the acquisition of (complex) syntax and morphology (Gass 2004; Gass et al. 2007; Mackey 2007).

Even if one may not agree with particulars of the IIO model, the super-theory of SLA (Block 2003), it would seem that the notion of Input is a central part of the epistemology of traditional SLA, the *sine qua non* of the field (Gass and Mackey 2007). Frequency of linguistic material in this input, then, will also have a role to play in traditional SLA; theories differ on the importance, though. Frequency in linguistics and L1 acquisition research has received a fair share of attention recently, especially in the volume edited by Bybee and Hopper (2001) where frequency as a phenomenon is studied in its own right; frequency, it is claimed, "bring[s] about form in language", and for the individual using and learning language, frequency "has a profound influence on the way language is broken up into chunks in memory storage" (Bybee and Hopper, 2001: 3). Frequency is also a major feature in linguistic description within the UBL framework in general; it is seen as the very aspect that blurs the distinction between language knowledge and language use, the argument being that

the way frequency shapes and is shaped by the linguistic system as this system is put to use "highlight[s] the dynamic interplay between language use and the speaker's linguistic system". (Barlow and Kemmer, 2000: xi). In other words, what is frequently employed and thus institutionalised in the speech community is also likely to be entrenched in the mind of the speakers (N. Ellis 2002; also Hoey 2007).

Issues of frequency have a role to play in my research papers as well; however, frequencies of expressions in a speech community, or even in terms of the input in the classroom which forms the social context of my data, are ignored. The frequency effects that I am concerned with pertain solely to the individual linguistic inventories. In terms of environmental influences of those emergent linguistic inventories that I investigate, I assume, along with N. Ellis (2002), that what is entrenched in the minds of my focal students is also likely to be frequent in the various environments in which my focal students navigate, be they in or out of class.

In a 2002 issue of *Studies in Second Language Acquisition*, N. Ellis (2002) forcefully argued in favour of giving aspects of frequency a more prominent role in SLA research. A number of well-known scholars were invited to write response papers to N. Ellis's paper, and while his thoughts were generally well-received they were also met with some opposition. It is not the purpose of the present section to go through all the responses to N. Ellis's paper; rather, the point is to put into perspective a field-internal inconsistency previously mentioned, namely that concerning predetermined learning trajectories, the individual as an individual engaged in active hypothesis-testing, and individual differences.

Noting that the question of frequency for SLA is unresolved, Gass and Mackey (2002), in their response paper, speculate that both token and type frequencies in interaction may play an important role, as language learners are engaged in on-line language use. They also argue, however, that in terms of universal developmental sequences, "acquisition appears to proceed at its own pace regardless of the frequency in the input" (Gass and Mackey 2002: 253). Tarone (2002: 291), in her response, seems to be stating exactly the opposite, as she worries that "Ellis's strong emphasis [on frequency in SLA] may lead us to think of [...] interlanguage as something that is [...] passively and unconsciously derived from input frequencies." While Ellis himself (2002) is careful not to underestimate the power of conscious noticing, in Schmidt's (2001) terms, in SLA – and further

consenting to the idea that "learners have to FIGURE language out" (Ellis 2004: 53; emphasis in original), it is noted with interest how Gass and Mackey find that the passive acquisition of structure as reflected in universal developmental sequences is uncontrolled by aspects of frequency, therefore questioning the importance of frequency, whereas Tarone is sceptical of ascribing too much importance to frequency at the expense of a view of the individual language learner as a cognitively active participant in L2 learning. In other words, Gass and Mackey view frequency as somewhat problematic because many aspects of L2 learning are automatic, whereas Tarone's view of frequency is that it is problematic because it implies a great deal of automaticity on behalf of the L2 learner; language learning, she argues, is largely a conscious activity (which, as noted, N. Ellis emphatically does not disagree with).

This field-internal discussion highlights very nicely an anomaly within SLA research referred to earlier, namely the bi-polar, almost self-contradictive, obsession with simultaneously identifying universal sequences and focussing on the L2 learner as a cognitively active hypothesis-testing individual. In fact, it now becomes clear that the only reasonable way out of this dilemma is to go with Pienemann and posit processing constraints on the learner combined with a predetermined learning trajectory. That way, the hypothesis-testing activities are naturally limited, both cognitively and linguistically, minimizing the possibility of hard-to-explain individual differences.

On a more epistemological note and irrespective of individual researchers' take on aspects of frequency, the input-interaction-output model is based on some assumptions that may or may not be empirically valid, but which have come to be truisms in the field at large. Part of the heated debate initiated by Firth and Wagner (1997), these include the notion of an explicit link between interaction and learning, the notion of the non-native speaker (NNS) as first and foremost a flawed communicator, and the notion of interaction between NNS and native speaker (NS) as a place in which ideas are singularly transferred from head to head with the primary purpose of promoting the development of the NNS' interlanguage grammar.

Looking first at the nature of NS-NNS interaction, the turn-taking machinery for everyday conversation as laid out by Conversation Analysts Sacks, Schegloff, and Jefferson (1974) and widely accepted as a robust finding in the social sciences has sometimes been argued not to work, or work only very problematically, for NS-NNS interaction (Long 1983). Instead, NSs have been

found to modify their language when interacting with NNSs. While such a notion of 'foreigner talk' is empirically valid and goes back to Ferguson (1975), it is not in opposition to general turn-taking mechanisms. In terms of Conversation Analysis (CA), this flies under the banner of 'recipient design' (Sacks and Schegloff 1979; quoted in Hutchby and Wooffitt 1998) which denotes that a speaker always designs his/her turns at talk in certain ways, depending on what he/she assumes to be existing mutual knowledge between him/her and the conversational partner. In CA, then, what turns-at-talk mean, or more importantly what they *do*, in interaction is defined by the participants (emic perspective). To IIO-researchers, who view interaction and input as separate entities (Long 1983), the turns-at-talk, deployed by NSs in NS-NSS interactions, are modified linguistic forms which, as seen from the viewpoint of the SLA researcher (etic perspective), are supposed to help propel NNS interlanguage grammar development forward. L2 learning, from this perspective, is measured solely in terms of increased grammar correctness, and interactions in which this grammar development supposedly takes place are investigated not because of their social interactional value but because they constitute a mere locus for interlanguage grammar development (Hall and Verplaetse 2000).

Thus, the notion of the NNS as a flawed communicator in need of improving formal mastery of the L2 combined with a view of NS-NNS interaction as modified to include negotiation for meaning in which the NNS is brought to *notice* (Schmidt 1990) what he/she is yet to incorporate into his/her interlanguage system, lead proponents of the IIO model, in Gass's (1998: 84) words to work "to understand what types of interaction might bring about what types of changes in linguistic knowledge". While they claim that the fertile relation between interaction and learning is a robust finding and clearly established, the fact of the matter is that, even though interaction may ultimately be the optimal locus for acquisition, a direct causal link between interaction and learning is extremely difficult to come by (Kasper 2004; Firth and Wagner 2007). This has even been acknowledged by Gass herself (Gass 2003). Given this lack of an explicit causality between interaction and learning, it might be more fruitful, as maintained by critiques of the model, to think of interaction and learning as two ends on a continuum; it is impossible to pinpoint exactly where one stops and the other begins. I explore precisely this territory in ESK1, ESK3, and ESK4.

As an example of this learning issue, below is a piece of data from Gass (2003: 235) / Mackey (1999: 558-9) (NS = native speaker; NNS = non-native speaker). As the authors explain, the extract is elicited from a picture-based 'spot the difference' task:

- NS: There's there's a a pair of reading glasses above the plant.  
NNS: A what?  
NS: Glasses reading glasses to see the newspaper?  
NNS: Glassi?  
NS: You wear them to see with, if you can't see. Reading glasses.  
NNS: Ahh ahh glasses glasses to read you say reading glasses.  
NS: Yeah.

At first, Gass (2003) notes that the NNS acknowledges (Gass's term; it is unclear what it means in this context) that the new word 'reading glasses' comes from the context, from the meaning negotiation specifically. She argues that this form of negotiation for meaning is particularly fruitful in terms of propelling the acquisition forward, as there is ample opportunity to pick up 'language' – in this case, the item 'reading glasses'. On a later page, however, Gass herself questions the generally posited causality link, 'uptake', between interaction and learning; a link which remains elusive. According to Kasper, Gass lists only three studies which confirm the link. Depending on how one reads Gass (2003), however, one might find more – or less than three. Below is a summary of some of the studies mentioned in Gass (2003), for a full account, please refer to the source; see also Mackey (1999) for a discussion<sup>92</sup>. The list shows the effect of interaction on issues of learning and production and as will be inferred, maximally two studies show an effect of interaction on learning (note: not an explicit link). The rest show differing degrees of impact of interaction on learning and production.

Study:	effect of interaction:
Gass and Varonis (1989)	positive (however, see below)
Gass and Varonis (1994)	on production, not learning
Mackey (1999)	leads to development; quantitatively increased participation leads to increased development

---

<sup>92</sup> For an overview of more recent studies, see Mackey (2007).



Sato (1986)	none
Loschky (1994)	inconclusive
R. Ellis et al. (1994)	positive on comprehension and learning of new words. Object of study was negation and word order; no results given on that.
Polio and Gass (1998)	positive on production.

For example, Gass and Varonis (1994) and Polio and Gass (1998) are listed as supporting the claim that interaction has an effect on L2 production but not necessarily on learning, which in itself is a strange dichotomy, and Mackey (1999) is listed as claiming a relationship between active involvement in interaction and enhanced development. To me, there is a rather big difference between claiming a relationship between active participation and linguistic development in the L2 and claiming an explicit causal link between interaction and learning. The former is compatible with a continuum between interaction and learning where the two are not kept apart, whereas the latter keeps them apart by inserting an important, researchable link between them.

A further problem, paraphrasing Kasper (2004), is the one of incipient learning. In the debate between psycholinguistic and sociolinguistic approaches to SLA, initiated by Firth and Wagner (1997), see below, one of the criticisms against the socially inclined researchers from mainstream SLA was that they really were not concerned with individual acquisition over long time (Gass 2004; Larsen-Freeman 2004). Also pointed out by researchers sympathetic to Firth and Wagner's (1997) proposals for a reconceptualisation of some established SLA concepts (e.g., He 2004; Mori 2007), they largely accepted this, acknowledging that longitudinal microanalysis was still in its infancy as research methodology, but this is now being developed (Brouwer and Wagner 2004; Hellermann 2006, 2007; Hellermann and Cole *forthc.*; Firth and Wagner 2007). Interestingly, Gass (2003) also discusses this as a problem for the IIO-model; that the data largely fail to show any long-term learning (acquisition) to have taken place. She refers to a piece of data in which a NNS mispronounces a word and then pronounces it correctly roughly 20 turns down the interactional lane. Gass and Varonis (1989) take this to be evidence that the item has been incorporated into the learner's grammar. It is not even certain, from the face of the data, that pronunciation was an issue between the two speakers in the interaction, let alone that any negotiation or learning took place, but even if some kind of attention was given to the sounds of the item in question and learning was

taking place, it cannot be argued, based on this extract alone, that anything but incipient learning was taking place.

Intriguingly, it is the conceptualization of acquisition as 'uptake', inherent in the IIO-model, that makes the model itself at best empirically problematic, at worst empirically impossible, resulting in the collapse of the whole construct. The idea that there is a rather precise point of uptake of linguistic material in the course of NNS-NS interaction is what separates 'interaction' and 'learning' into two sealed-off units. However, "uptake is notoriously unreliable as a means of establishing that something has been learned" (Larsen-Freeman 2004: 606), and interactional data simply never (or very close to never; Kasper 2004) confirm the point of uptake of any linguistic material. Operationalizing acquisition differently, e.g., by putting it on a scale at whose other end we might position 'social interaction', the model might not be that far off the point; in fact, the current formulation of a "connection between interaction and learning" (Gass and Mackey 2007: 176) is a statement few would contest. However, critiques of the IIO-model, usually more socially oriented scholars, are likely to maintain the indivisibility of the two rather than one explicitly causing the other, with the elusive notion of 'uptake' floating in-between. Whereas Gass and Mackey's approach may thus be termed a "learning-in-interaction" approach, the socially inclined researchers are more likely to be speaking of "learning-as-interaction" exactly because the point of uptake is elusive. I position myself in ESK1, ESK2, ESK3, and ESK4 as an ally of the latter grouping, as my data have led me to abandon the 'uptake' metaphor.

In other words, few would object to the notion that interaction is a hugely important aspect of L2 learning; the point of contention is the exact relationship between the two. Viewing interaction as integral to learning is not the prerogative of the IIO-model; it is given paramount importance in the present research as well as in numerous other approaches existing in the vibrant field of SLA today, such as CA for SLA (e.g. Firth and Wagner 1997, 2007; Markee 2000; Brouwer and Wagner 2004; Mondada and Pekarek-Doehler 2004; Hellermann 2007, 2008); socio-cultural and socio-cognitive approaches to SLA (e.g. Lantolf 2000, 2005; Atkinson 2002; Hall 2004; Watson-Gegeo 2004; Lantolf and Thorne 2006; Atkinson et al. 2007); and second language socialization studies (e.g., Kanagy 1999; Zuengler and Cole 2005; Hellermann 2006). What separates the strands is the very nature of interaction. Socially oriented scholars usually view interaction as a site of organized social action in which people *do* things, accomplishing activities sequentially, even down to turns-at-talk,

often in a co-constructed manner (Lerner 1991). The IIO-model, on the other hand, assumes interaction to be mainly about transmitting information from one individual mind to another. Combined with the single-minded focus in the IIO-model on NNSs as flawed communicators and interaction as their chance to test hypotheses, negotiate for meaning, and 'take up' linguistic items from the input, the model remains an hypothesis which is yet to be empirically validated – and yet its foundations have been accepted in traditional SLA circles at large as robust and true. So to some, perhaps, IIO is *the* super-theory of SLA; I just fail to see its validity in my data.

Intriguingly, Long (2007), in his recent lengthy blatant criticism of Firth and Wagner (1997) – and others who wither agreed with Firth and Wagner or just advocated the notion of theory proliferation in SLA – argues that socially inclined researchers need to bring forth evidence of their claims, rather than merely stating their epistemological preferences. Obviously, there is nothing wrong with what Long justifiably demands; the problem with Long's discussion is that he either ignores or is unaware of the accumulating facts: evidence virtually abounds that social interaction – not mere negotiation for meaning – has consequences for learning. Apart from the references above and the present research as displayed in ESK1, ESK3, and ESK4, Larsen-Freeman (2004, 2006, 2007) is increasingly aware of the need for a view of L2 development that is more balanced in terms of the social-cognitive discussion. The evidence simply points in that direction. Acknowledging this is just a matter of getting acquainted with the research. Long's own model, IIO, on the other hand, remains to be empirically validated as far as the 'uptake' metaphor is concerned.

What I hope to have shown so far is the birth and development of the individual-mentalist orientation of the traditional field of SLA. In the process of thus describing the mechanisms which have shaped the field, I have also implied but never fully articulated the self-entrenchment, as it were, of the field. This tendency within SLA to protect its outer boundaries against agitators with a different agenda became visible for all in 1997-8 in a series of articles in *The Modern Language Journal*, starting with the seminal piece by Firth and Wagner (1997). The repercussions of this debate between the two sides, which are sometimes referred to as psycholinguistic and sociolinguistic, respectively (Long 1997; Atkinson 2002; Block 2003; Markee and Kasper 2004; Firth and Wagner 2007), still reverberate throughout the field. Firth and Wagner (1998) argue that there is no clear dividing line between psycholinguistics and sociolinguistics, and today, prominent

SLA scholars, previously psycholinguistic in orientation, tend to view the intersection of the two approaches as the most fruitful point of research for future SLA (e.g., Larsen-Freeman 2007).

### 8.3.2 Tracing the social-cognitive split.

In the discussion below of various dissident tendencies within the broader field of L2 studies, I will try to portray the debate as seen from my perspective. It will not be a neutral survey; rather, I will state my affiliations when necessary and relevant for the point I will be making in favour of embracing what I term a new SLA eclecticism. As I remarked earlier, 'environmentalist' views have historically been shunted to the backwaters of modern SLA; behaviourism one might argue, was *all* environmentalist, but with its stimulus-response psychology it was abandoned as Chomsky managed to convince the world of linguistics that recursive rules of syntax, ensuring an infinite ability to embed clauses in clauses, guaranteeing syntactic creativity forever, constituted the stuff of language and was thus at the heart of what L2 learners had to acquire. In retrospect, then, the germ of the division of the field into a sociolinguistic and a psycholinguistic branch (Markee and Kasper 2004; Firth and Wagner 2007) can be traced in SLA's history; through the 90s (Long 1990; Larsen-Freeman 1991) and all the way back to Hymes, essentially, who, as we also noted, ran into closed doors (or selectively open windows, to remain more faithful to his own choice of words) with his conflation of language knowledge and language use. Summing up before moving on to a more detailed discussion of the division of the field, aspects of sociality have historically been kept outside probably due to the bias of the field, tuned as it was to individualistic, cognitivist, syntactocentric foci.

Firth and Wagner (1997) sparked off the debate as they invited for certain changes within the field at large<sup>93</sup>. They argued that the individualistic and mechanistic view of discourse and communication in mainstream SLA had obviated insight into the language use of S(econd)/F(oreign)L(anguage) speakers. Instead they endeavoured to integrate the social and cognitive dimensions of S/FL use and acquisition by rejecting Chomskyan dualisms and the

---

<sup>93</sup> A detailed meta-debate on the 1997-8 debate can be found in *the Modern Language Journal*, 91, 5 (2007). As pointed out by Lafford (2007) and Larsen-Freeman (2007) the social-cognitive debate in SLA predates Firth and Wagner (1997). Firth and Wagner's paper did, however, turn the issue into a prominent one within the field, resulting in further discussion on the dualisms of applied linguistics (N. Ellis and Larsen-Freeman 2006). I maintain that the discussion between Gregg and Tarone / R. Ellis in the early 90s on variable competence, cf. previous section, is not parallel to the current discussion to the extent that the contenders in the debate did not take issue with the basic question of the nature of Interlanguage (the target-language perspective, Larsen-Freeman 2007). Therefore, that discussion will not be referenced here.

dominating view of interlanguage as movement towards a fixed end-point as increasingly complex structures are acquired. While IIO-researchers saw no need to follow this lead, others have taken the encouragement to heart (e.g. Atkinson 2002). Recently, psycholinguistically oriented scholars, experienced researchers in the field, N. Ellis and Larsen-Freeman (2006) and de Bot et al. (2007), have argued in favour of viewing both language and language learning as simultaneously social and cognitive. We shall get back to these approaches.

Researchers who agreed with Firth and Wagner that the hitherto practice of marginalizing or ignoring contextual dimensions of language in SLA was deeply problematic went on to conduct research that seemed exclusively to deal with local interactional aspects of SLA, rather than developmental ones (such as Mondada and Pekarek-Doehler 2004). It remains a challenge to give a full interactional account of developmental issues in SLA (Brouwer and Wagner 2004), which was also acknowledged by Kasper (2004), who stated that her application of CA to SLA had said nothing about acquisition. As mentioned above, however, a practise of doing longitudinal CA for SLA is now being developed.

In terms of Firth and Wagner's (1997) argument that Chomskyan thinking, or at least psycholinguistic assumptions prevalent at the time of birth of the field, led to a preoccupation with individualistic and mentalistic foci in SLA at the cost of social and contextual aspects of language, I agree. However, new theoretical approaches to language and linguistics have emerged that are fundamentally un-Chomskyan but still focus on language knowledge of individuals. Such an approach is UBL as outlined in chapter 6 and applied to L2 developmental issues in the five research papers. These papers make it clear that some aspects of language knowledge are generally portable by the individual L2 learner – but they also suggest that some aspects of language knowledge are highly situated. This seemingly empirical fact of the situated nature of aspects of linguistic knowledge is extremely important in terms of researching L2 development, because such development rests on numerous instances of situated language. The framework of UBL is particularly apt at describing and analysing the similarities and differences between situated and portable aspects of language knowledge because they are seen as matters of degree along a continuum of schematicity; the more concrete and lexically specific the language, the more situated its nature. Furthermore, the UBL framework with its assumption that language learning is a matter of schematising in ontogenesis that which is lexically specific seems an empirically valid model of

L2 learning as explored in the five research papers. So, individuality is not necessarily a Chomskyan problem in latter-day SLA, let alone a problem at all. Therefore, the major problem in SLA, I suggest, is psycholinguistic compartmentalization of linguistic knowledge and the ideas following it that (S)LA is about learning abstract rules of syntax. As suggested in EC and ESK4, L2 learners seem to learn formal regularities as construction-specific patterning which is not reducible to maximally general rules used for combining the items in the constructions; rather, linguistic knowledge is an experiential inventory of such patterns.

The problem thus outlined is one of operationalizing the nature of language knowledge, both in terms of the transitional characteristic of 'learner' language and the dynamic and distributed nature of 'mature' language knowledge. In terms of traditional SLA, as discussed in the previous section, interlanguage theory assumes pre-determined stages of morpho-syntactic acquisition, based partially on linguistic and partially on psycholinguistic constraints, en route to near-native mastery of target-language structure. Some research in this interlanguage vein has been explicitly Chomskyan either because it has investigated the relevance of UG for SLA (e.g., Clahsen and Muyskens 1986; White 2007) or because it has investigated Chomskyan aspects of grammar such as derivation of surface word order (Meisel et al. 1981) or acquisitional order of Determiner Phrases (Zobl and Liceras 1995). More often than not, however, there is kinship with Chomsky in terms of compartmentalization given the lopsided focus on morpho-syntax, most recently stated in Gass et al. (2007).

Another problem pertaining to the issue of language knowledge discussed by Firth and Wagner (1997) goes back to the old baseline discussion already present in early interlanguage writings (Corder 1967; Selinker 1972). While I do not agree with Firth and Wagner's (1997) equation that a view of L2 users as learners in a state of transition automatically entails a focus on language skills and competences as underdeveloped, I think they rightfully point out that the interlanguage notion of transitional phases as systematic with native speaker competence viewed as stable is problematic. The clinically sober and fairly balanced observation that L2 users develop their abilities to interact in the target language is, I would argue, the *raison d'être* of a field that is concerned with L2 learning. The question is how we operationalise 'abilities to interact', 'target language', and 'learning'. Furthermore, maintaining a view of language knowledge as fundamentally transitional, as Firth and Wagner (1998) elaborate, automatically entails that everybody is always and everywhere

in a state of transition. Transition does not necessarily entail underdevelopment, let alone a lopsided focus on it. A case in point here is ESK4 which shows situational development of a learner pattern in task expansions in the classroom; a learner pattern which from a target language perspective would not be considered a token of 'acquisition', but which was found to be dominant in this particular learner's biography to such a degree that it had to be taken seriously on its own terms. The learning of the pattern could even be said to be item-based as it was found to be dependent on a recurrence of one particular instantiation, namely *you no write*.

This 'baseline problem' also surfaces in Kasper (1997), who, in her response article to Firth and Wagner, supports the need for measuring against a standard in SLA (also Poullisse 1997), without defining in any terms how she conceptualises such a standard. Referring to Gregg's explicitly Chomskyan work in SLA, she does point out, however, that one does not have to buy into the Chomskyan notion of Competence in order to acknowledge that SLA is about 'acquisition'. I fundamentally agree with this position; I have no problem with the 'acquisition metaphor' (Sfard 1998) for learning; in fact, it is implicitly assumed in EC and I explicitly argue in favour of it in ESK3. I do however find it very problematic that Kasper does not outline her views on what constitutes (mature) language knowledge at this stage. Precisely because Kasper does not go into any discussion of the nature of language knowledge, either in terms of interlanguage or in terms of native language knowledge the reader does not know if she subscribes to, say, a maximalistic model such as UBL or a more compartmentalized view of language knowledge; the latter, obviously, would be parallel to prevailing views in mainstream psycholinguistic SLA. What is further disconcerting in terms of the prospects of applying a different linguistic apparatus than formalism is that Kasper seems to retain the use-acquisition dichotomy, next-in-kin to Chomsky's own competence-performance distinction, as well as the idea related to it that second language acquisition is describable, its essential problems approachable, as an incremental interlanguage journey through stages of increasing structural complexity. The main problem, she argues, with Firth and Wagner's paper is that they do not address these issues. The fact is, however, that they do; they state their predisposition against that view of language learning. It would therefore seem that not only does Kasper here subscribe precisely to the kind of syntactocentrism that much interlanguage work revolves around in its quest for answers pertaining to how L2 learners acquire the structure of the target language, she implicitly argues that such endeavours define SLA research.

And therein lies the rub, it seems; going back for a moment to the SLA historicity reiterated above, we noted with e.g., Long (1990) and Larsen-Freeman (1991) that there is general consensus that the modern field of SLA came into being at a point when behaviourism was going out of style and it was becoming legitimate if not desirable, most prominently with Chomsky (1965), to make inquiries about the workings of the human cognitive system. The scope of what SLA theory and research must explain in order to be accepted into mainstream research circles has not fundamentally changed since then (Lafford 2007). Firth and Wagner (1997) also make this point which is then fundamentally misunderstood by Long (1997) in a lengthy footnote to his response piece. He claims that Firth and Wagner in service of their argument superimposed UG on all psycholinguistic SLA research; in fact, they did no such thing. They argue that Chomskyan conceptions of what it means to use, know, and acquire a language have been pervasive in SLA, wiping aside all rivalling linguistic schools of thought. They quote one of the founding fathers of modern SLA, Corder, who reiterated Chomskyan preconceptions about the individuality of language and SLA's concern with underlying psychological mechanisms of language use. Their quote (Firth and Wagner 1997: 287) is very similar in nature to Chomsky's (1965) programmatic statement in *Aspects of the Theory of Syntax*, already quoted earlier, that linguistic research should be concerned with finding a mental reality underlying linguistic behaviour. This Chomskyan tendency in interlanguage studies, so obvious in the shared epistemologies between the two, can hardly be said to be controversial.

In the debate on sociolinguistic approaches to SLA and how they may generally contribute to SLA, it is usually inferred that use is for the sociolinguists, acquisition for the psycholinguists. However one approaches this problem, then, entrenched SLA researchers may always dichotomise the issue, making the argument that the more socially oriented researchers can do their stuff, while the traditional researchers can continue to do their stuff. A case in point is Gass's response article (Gass 1998), which explicitly spells out the dichotomous relationship between use and knowledge, and the idea of language knowledge as autonomous and modular. In such a climate it is futile to speak of making more room for finer-grained interaction analyses. As pointed out by Hall and Verplaetse (2000), the research tradition represented by Gass is rarely interested in the social situatedness of interaction. With this in mind, if one wishes to redirect SLA, the argument needs to be won on the terms of the opponent. This means that as a first step in the debate, rather than quarrelling over the role and nature of interaction, an issue too easily dismissed by invoking an army of dualisms, we



need to theoretically validate and empirically substantiate the idea that language knowledge is fundamentally dynamic and transitory (Firth and Wagner 2007). Similar ideas are beginning to gain foothold in more traditional cognitively and individually oriented SLA circles where it is argued that a static linguistic theory is fundamentally misguided as starting point for the analyses of something that is essentially dynamic and, indeed, chaotic (Larsen-Freeman 2002; 2006). In other words, interlanguage theory in its original guise is gradually losing strength, it would seem; its notions of endpoints of learning increasingly degraded with emergentist ideas in the ascendancy (Lantolf 2005; N. Ellis and Larsen-Freeman 2006; MacWhinney 2006). Such a climate is friendlier to ideas pertaining to the relevance of going more local, as it were, in terms of interaction analyses. In fact, as briefly mentioned in ESK2, linguistic theoreticians in UBL, Fried and Östman (2005) are beginning to realise the compatibility of methods and approaches in construction grammar, interactional linguistics (e.g., Ford 2004), and Conversation Analysis, arguing in favour of compatibility between the turn-constructive units of Conversation Analysis (e.g., Sacks et al. 1974; Schegloff 1996) and the form-meaning pairings of UBL. While the compatibility may not be straightforward, primarily because of fundamental epistemological differences among the approaches, such ideas are in alignment with the SLA eclecticism advocated here and constitute a point of interest for future research.

Gass (1998) goes even further in terms of the dualisms upheld. As a result of her own main research objective, to understand what aspects of interaction determine interlanguage development, she maintains a strict division between language used and acts of communication. Also commented on in Firth and Wagner (1998), this epitomises the apples and oranges of the title of her paper. Gass's tenet underlines SLA's reluctance towards integrating interactional analyses in research in L2 learning and development. Such aspects may be discarded in Gass's framework (Gass 1988) because the researcher is only interested in the input and how that might potentially drive interlanguage forward as the learner notices it, processes it, and takes parts of it up – all in order to restructure the interlanguage. Such a research interest requires a willingness to decontextualize both learning process and the subject matter to be learnt, i.e., language.

So unless we manage not only to argue persuasively but to show empirically that neither learning processes nor language knowledge 'happen' in a social vacuum, proponents of the IIO-model may all too easily dismiss criticisms due to epistemological differences. Instead, the very act of

communication, in UBL terms, what goes on in *the usage event*, done away with by Gass, needs to be reinstated as the object of interest to the researcher as it is the object to be learnt by the L2 learner. Learning such acts of communication requires participating in real interaction with other real people, and it requires the learning of interactionally useful linguistic units, namely MWEs, utterance schemas, and, at its most abstract and perhaps at most advanced levels of learning, constructional schemas. Fine, if Gass insists that communication and language are two different phenomena, so be it. But the field of SLA will benefit greatly from denouncing such an approach. Instead language and communication must be viewed as inseparable; you cannot have one without the other, and so L2 learning is about language (knowledge) and communication at the same time. The contents of the linguistic inventory are shaped by linguistic experience, which in turn develops in use, i.e., communication, in *usage events*.

Involved in the experiential learning of such an inventory as I explore it in the five research papers, is the learning of linguistic items, patterns, that may emerge as abstract schemas if they are used frequently enough in the speech / language learning community. However, some patterns may not be describable as more than locally contingent ones, applicable in a range of situations, and partially consisting of fixed items. Some patterns may even be purely fixed items made relevant in recurring social situation (formulas). This is an empirical fact (e.g. Kanagy 1999; Eskildsen and Wagner 2007) that calls for an integration of interactional phenomena into explanatory aspects of SLA at large. But all linguistic patterns, from fixed patterns to general schemas, are equal parts of the inventory, and the former do not by definition, in ontogenesis, turn into the latter. There is no a priori given acquisition of increasingly abstract structures. In other words, the task is to theoretically argue and empirically substantiate the opposite ideas; this would answer Long's (2007) criticisms that what we might term the social viewpoint is nothing but epistemological postulates. I think the development of the themes as explored in the five research papers, going from a fairly decontextualized analysis of L2 development towards an increasingly locally contextualized account of L2 learning marks an attempt in this direction. I take sides with Larsen-Freeman (2007) that the challenger's view (i.e., Firth and Wagner's proposal for a reconceptualisation of some tenets of SLA) must be reckoned with; and I think that the five research papers in this anthology support the idea that individual linguistic development may be investigated as such but most profitably, if not exclusively, with descriptive and analytical recourse to the social environments in which development takes place. It is simply wrong to claim (e.g., Long 2007) that social setting does not

inflict upon 'competence'. Precisely because language knowledge, or at least important aspects of it, have been shown to be socially situated, social environments must, by implication, matter to learning.

In hindsight, then, it appears that the debate initiated by Firth and Wagner (1997) on a redirection of the field of SLA was undertaken partly on the wrong terms. Debating the nature of interaction on SLA has been futile; not in the sense that the debate left no mark on L2 research, but in the sense that the field at large has seemed immune to the arguments concerning a broader L2 database and a re-sharpened focus on interaction *per se*. Such matters are too easily shunted to the muddy backwaters of language use by proponents of the IIO-model. Instead, we must make the point that it is fundamentally wrong seen from a UBL perspective to modularize language and separate use and knowledge – and that the ensuing separation in SLA of interaction/use and acquisition was equally wrong. Learning is a cognitive issue (Gass 1998), yes, but not exclusively so. Traditional SLA is in danger of being isolated on this one, with socially anchored theories of learning vastly growing, including language socialization for SLA (Kanagy 1999; Hellermann 2006), Socio-Cultural Theory (Lantolf 2000; 2005); CA for SLA (MLJ 2004); even psycholinguistic researchers N. Ellis and Larsen-Freeman (2006), as already mentioned, have recently accepted that both language and learning are socially contingent. We are, it seems, in a process of changing the conceptual kit of SLA (Rampton 1997) as we increasingly integrate interaction and acquisition, ask questions about how the individual linguistic inventory, the portable language ability, that arises out of interactional contexts is made relevant and recontextualized in new comparable interactional contexts, and about the nature of the predominant unit in that inventory. This change allows us to refocus the field in terms of the prevalent assumptions about both interaction and linguistic knowledge, and begin to look outside the sphere of psycholinguistics for explanatory forces in L2 learning. Gass (1998) does acknowledge that a view of language as socially constructed does not necessarily entail the impossibility of viewing language as something abstract residing in the individual. Arguably she is right, it does not. Setting aside for now the position that the abstract nature of individual language knowledge is yet to be empirically validated, we can safely conclude that Gass's own position, maintaining a strictly mentalist view of language knowledge has prevented her from accepting the fact that language is, also, a social phenomenon, explanation in SLA still to be found in the psycholinguistic workings of the IIO model (Gass 2003).

Poulisse (1997) also frames her response to Firth and Wagner around the idea that because acquisition is generally thought to be cognitive in nature it is no wonder that SLA is cognitive. If one is studying psycholinguistic processes, one should relate one's findings to psycholinguistic theories, so the argument goes. But that is not what is being done in SLA; rather, in the field per se people study SLA and relate findings to psycholinguistics. They do not seek explanations elsewhere. The environment, in the form of input, is important for the psycholinguistic processes of learning a L2, but it never seems to represent more than that; it is the presupposed locus for acquisition. Even if one assumes with Kasper (1997) that one does not need to buy into the UG package to do psycholinguistic SLA, her own views on what language is and thus what constitutes language knowledge remain elusive at best. Like Kasper, and essentially also Poulisse and Long, however, I too am fairly comfortable with the cognitivist position promoted in much traditional SLA research. There is nothing wrong with cognitivism, there is nothing wrong with individualism, either, it is the compartmentalized linguistic apparatus usually invoked that presents the biggest problem with its lopsided focus on some syntactocentric mental reality underlying behaviour, excluding any aspects of the environment for having explanatory power in L2 learning. This also seems to be the essential thinking in Vygotskian Sociocultural theory (on which more below) as it is put forward in the response article by Hall (1997: 301) who talks about individual development as "movement from the social to the psychological", while problematizing the modular and decontextualized view of language use, knowledge and acquisition.

In her response article to Firth and Wagner (1997), Gass (1998) argues in favour of a few established key notions in traditional SLA, among them being learners as the sine qua non of the field, the importance of deficiencies in learner language, and, related to both previous points, the cardinal question for SLA concerning the nature of learner systems. First, nobody, not even Firth and Wagner themselves, would contend with the notion of learner. It is, after all, used widely in various kinds of classroom research, including practice-oriented studies (Hellermann 2006, 2007; Hellermann and Cole *forthc.*). Rather, Firth and Wagner attempted to broaden the notion, to have the SLA database include L2 users that were not necessarily primarily formal classroom 'learners'. In line with their own argument (Firth and Wagner 1998) that everybody's version of a language is in a state of constant change, everybody might qualify as a learner. The problem lies in the lopsided focus on deficiencies in learner language. Gass argues, probably rightly so, that this focus is not on the deficiencies themselves; rather, the deficiency focus abounds because of the underlying

interlanguage assumptions that 1) learner language is always systematic with descriptive recourse to rules of syntax, and 2) the learner traverses a path of approximating the perfect, rounded, native target language, which necessarily implies increasingly fewer deficiencies in the interlanguage. The idea that learning does not follow interlanguage systematicity, and that learning does not necessarily entail movement towards structural completion or native-like systematicity, has only recently moved from status of heresy to authority in cognitivist SLA as witnessed by the emergentist proposal by N. Ellis and Larsen-Freeman (2006). The present research as displayed in the five research papers seek to explore such emergentist ideas.

The SLA establishment, then, did not respond favourably to Firth and Wagner's call for a reconceptualisation of some central aspects of research within the paradigm. Situating SLA more in terms of social performance was not a move the field at large was interested in. Though they did find kinship in places (e.g. Liddicoat 1997; Hall 1997; Markee 2000; Atkinson 2002; Block 2003; Markee and Kasper 2004), their support was limited to people who already took a social perspective on issues pertaining to language knowledge and learning; as Larsen-Freeman (2007) notes, the social agenda, generically speaking, was not new to the field at the arrival of Firth and Wagner's nonetheless influential piece. Even though Lafford (2007) comments on the equivocally seminal status of Firth and Wagner (1997) few would dispute its importance. Gass et al. (2007) even go so far as to say that the article did the field a disservice, arguing that it has prevented researchers from engaging in fruitful collaboration due to its confrontational nature. Perhaps Gass et al. are right, perhaps not; perhaps the result that they mention, that the article prevented fruitful collaboration, lay as much in aggravated responses from mainstream psycholinguistic researchers. It is implied by the term *confrontational* (used by Gass et al. 2007) that two camps are in battle; in any case, that was the result: SLA was divided, roughly, into two camps, a psycholinguistic and a sociolinguistic one (Markee and Kasper 2004; Firth and Wagner 2007). The former would stick to its established views on language and language learning, the focal point being exclusively individual and mentalistic with no room for a consideration of social performance, whereas the latter would go on to develop new ways of investigating L2 learning as a social accomplishment in a shared community of practice. Only very recently has there been a tendency for the field to approach an operationalization of a new form of SLA eclecticism parallel to the one advocated here, especially in Larsen-Freeman's recent work (Larsen-Freeman 2006, 2007; N. Ellis and Larsen-Freeman 2006).

### 8.3.3 Second language acquisition or second language participation?

Having been concerned primarily with the psycholinguistic tradition in SLA, most prominently reified by the IIO-model, and the challenge made against it by Firth and Wagner (1997), it is now time to turn toward the more socially oriented researchers; those who took sides with Firth and Wagner. This will eventually assist in placing the present research in a framework of its own somewhere in-between the social and psychological-mentalist extreme poles. As recently noted by N. Ellis and Larsen-Freeman (2006) and Larsen-Freeman (2007), the field of applied linguistics seems to be permeated with dualisms whose conceptualizations all have bearings on how we approach the study of how people learn a L2. The field spilt recounted in the previous section implied taking sides on a number of decisive issues, each framed around the (not too fertile) ground of such dualisms. In this section, I embark on a discussion of this tendency to cultivate the L2 debate in terms of dichotomies rather than continua. Whereas N. Ellis and Larsen-Freeman list 20 such dualisms, including the Krashenite acquisition-learning and the consciousness-related implicit-explicit learning dichotomies, I choose here to focus on what I believe are the 6 most research-defining ones. In alignment with N. Ellis and Larsen-Freeman, I include competence vs. performance, learning vs. use, and mind vs. society. Besides those, however, I will discuss three other dualisms, these being acquisition vs. participation (also included in Larsen-Freeman 2007), etic vs. emic perspective on L2 data, and one which I have termed locality vs. generality. Table 1 below summarizes those 6 dualisms.

Related to Chomsky's competence-performance distinction, the use-acquisition distinction has long been a matter of debate in SLA (e.g., Firth & Wagner, 1998, 2007). However, lesser dichotomies, or at least dichotomies that more rarely materialize in SLA debates, also abound, including one between sociality and individuality and one between participant-relevant locally contingent interactional phenomena and generally (context-independent) applicable linguistic resources that transcend the moment (paralleled in the discussion of etic vs. emic categories of analysis as it is found in Firth and Wagner 1997; Markee and Kasper 2004). Last, but not least, we have one between two views of learning, as captured in the metaphors 'acquisition' and 'participation' (Sfard 1998). This section undertakes a theoretical exploration of these dualisms and will, in alignment with the data displayed in the five research papers, ultimately argue that in order to gain a clearer picture of development and change in longitudinal SLA studies, maintaining strict boundaries in terms of the dualisms mentioned must be avoided. In other words, the present mission is one of

diplomacy; gaps need bridging rather than further digging. This should not be confused with an agenda which has as its ultimate goal to unify the field. Heterogeneity is a strong point, perhaps a necessity (Block 1996).

First, however, let us have a look at the controversies from a relatively objective standpoint. I think it is safe to say that there are certain tendencies in the research that certain convictions regarding these dualisms go together. In other words, the parts of the dualisms form each a paradigm, as it were:

Socially oriented L2 studies	Psycholinguistic SLA
Use	Acquisition
Performance	Competence
Sociality	Individuality
Locality	Generality
Participation	Acquisition
Emic perspective	Etic perspective

*Table 1: delineating the field split.*

Table 1 depicts what has over the years become a fairly robust and well-maintained demarcation line in the field of SLA. Thus, not newsworthy in and of itself, this sentiment of a field division is still strongly voiced in the debate, recently in Markee and Kasper (2004), Zuengler and Miller (2006) and a special issue of the *Modern Language Journal* (Dec. 2007) on the impact of Firth and Wagner's 1997 piece. There is thus a tendency in the research that psycholinguistic SLA placed its focus on the notions on the right-hand side; in Block's (2003) words these approaches are most prominently represented by input-interaction-output hypotheses, promulgated by researchers such as Susan Gass and Michael Long. The research presented in the five research papers here, with its point of departure in UBL, focuses on individual linguistic inventories. Thus, this research has also been massively interested in individual ontogenesis; however, the trend in this research has pointed towards an increased interest in performance-driven linguistic inventories as well as an empirically reasoned need to more locally and interactionally contextualize the notion of the individual linguistic inventory. This movement towards an increasing focus on social context is evident in the development from ESK1 to ESK4. Other approaches – to some extent informed by the seminal

article by Firth and Wagner (1997) – to the study of what takes place when people learn a second language, e.g. CA<sup>94</sup> for SLA (*Modern Language Journal*, vol. 88, 4), traditionally placed themselves programmatically, and felt empirically substantiated accordingly, among the notions on the left in the table above.

It has sometimes been debated whether Firth and Wagner (1997) called for a redirection of the field of SLA per se or whether they encouraged it to loosen its strict mentalistic predispositions (Larsen-Freeman 2007). The following discussion will take that issue as its starting point and begin with approaches and ideas that have been put forward in the interest of clarifying some conflicts pertaining to the dualisms above. Four such approaches are *language socialization studies in SLA* (Watson-Gegeo and Nielsen 2003; Watson-Gegeo 2004; Zuengler and Cole 2005), *the socio-cultural approach* (e.g., Lantolf, 2000, 2005, 2007; Lantolf and Thorne 2006), *the sociocognitive approach* (Atkinson 2002; Atkinson et al. 2007), and CA for SLA (e.g., Markee and Kasper 2004; Mondada and Pekarek-Doehler 2004; Hellermann 2007; Firth and Wagner 2007). All have, to some extent, tried to bridge some of the gaps presented here but I find myself in sympathy with aspects of them rather than completely aligning with them. The point of the following is to briefly show how.

Researchers in language socialization studies in SLA have made the case for a redirection and redefinition of the field (Watson-Gegeo 2004). Building to a large extent on the work of Schieffelin and Ochs (1986), they set up a dichotomous relationship between language socialization and language acquisition, the former being about becoming competent members of a given social group (a process in which unspecified 'language' plays a role), and the latter, considered largely irrelevant in language socialization studies, being about strictly linguistic competence at different developmental points. Empirically, their investigations often concern interactional routines because they "serve as good locations for observing how socialization takes place." (Zuengler and Cole 2005: 304). One example of this is Kanagy (1999) who found a correlation between social and linguistic routines in her study of children being socialized into Japanese L2. Another example is Hellerman (2006), who shows empirically how classroom learners of English L2, in co-constructed dyadic encounters, develop their interactional competence as they are socialized into recurrent classroom practices, thereby becoming increasingly fuller members of the classroom community of practice.

---

<sup>94</sup> CA stands not for Contrastive Analysis but for Conversation Analysis.



While it is clear that such language socialization research makes tremendous contributions to SLA on its own terms, it is equally clear that it retains a fairly sharp distinction between learning as 'doing' or 'becoming' along the lines of the participation metaphor, rather than learning as 'getting' or 'accumulating' along the acquisition metaphor. Furthermore, and in relation to this, language knowledge seems to be captured in purely social terms. Since the interest of the present research as presented in the five research papers lies in exploring the possibilities of eclecticism, such clear stances on the dualisms prevent the socialization approach from taking center-stage here. I would argue that in order to challenge their own assumptions about language and learning, language socialization researchers need to go beyond the predictability of recurring social practices and routines. I do not doubt for a second that language learning is about social routines; I do doubt, however, that this is the full story. It must be an empirical issue to establish whether or not such relatively stable patterns of participation actually constitute the bulk of social situations in which language users engage.

Lantolf's Vygotskian socio-cultural theory (SCT) and the present research, as mentioned in ESK1, share the adherence to the emergentist idea that language knowledge is dynamic and that linguistic interaction is a constant source of renewal of the linguistic repertoire of language users (Lantolf and Thorne 2006). This is where straightforward compatibility ends, however. Essentially a theory of what cognition is and how it works, SCT primarily concerns how learning occurs as an 'internalization' process, describing the mechanisms which enable humans to control their higher mental abilities, including language. In ontogenesis, everything is experienced twice, as it were, first socially and then intra-psychologically, popularized as the tenet 'what you can do with others today, you can do alone tomorrow'. Learning, taking place in a person's zone of proximal development (ZPD), never happens in a social vacuum; "social interaction is the source of mental development" (Lantolf 2005: 342). For the present purposes, then, SCT seems a rather inflexible construct, considering its insistence on viewing learning as internalization in the ZPD, which is not easily positioned in terms of the learning metaphors of *participation* and *acquisition* explored here. Lantolf (2005) himself expresses skepticism toward the acquisition metaphor to the extent that it denotes sequentiality and predictability in learning processes, but a closer investigation of the construct invites for the interpretation that learning in SCT seems to be a very tangible process reminiscent of 'uptake' (a chosen term also in Lantolf 2005) as traditionally envisioned in input-

interaction hypotheses in psycholinguistic SLA. In other words, the very notion of 'internalization' seems to necessitate a very distinct demarcation line between interaction and acquisition, presupposing a rather clear causality aspect between 'use' and 'learning', the assumption seemingly being that the stuff of learning, in this case linguistic items, filters through from sociality to individuality as learning progresses. Curiously, this is indistinguishable from the traditional viewpoint prevalent in various psycholinguistic interaction hypotheses, as they have been proposed in various guises in psycholinguistic SLA by "authors ranging from Peter Skehan to Rod Ellis to Susan Gass" (Block 2003: 92); items encountered in the *input* are subject to *attention* or *noticing* (Schmitt 2001) by the learner's cognitive processing system in order to become incorporated into the competence system which is then *restructured*. Thus, at best 'internalization' is merely a confirmation of the prevailing operationalization of acquisition, and at worst it is an a priori claim that, in spite of its longevity, has never been empirically substantiated; as both Kasper (2004) and Firth and Wagner (2007) tell us, the proverbial uptake causality between interacting and learning is notoriously hard to come by. Evidence that a given piece of linguistic material is picked up by a learner because of attention to it in interaction is extremely hard to come by, as already mentioned. This strict view of causality is probably too simplistic to account for something as complex and seemingly chaotic (Larsen-Freeman 2002, 2004; Firth and Wagner 2007) as L2 learning. In other words, SCT is problematic for the present purposes because it seems to retain the traditional chasm between 'learning' and 'interacting'.

The sociocognitive approach (Atkinson 2002) seems like a promising way ahead for empirically driven, interactionally situated L2 learning research, its insistence to break with Cartesian traditions finding sympathy here, especially the inseparability of acquisition and interaction. However, because of its preliminary nature (Atkinson et al. 2007), some key issues in this line of work are still elusive, which makes it impracticable for the present purposes. It is unclear, for instance, how its proponents view the nature of linguistic knowledge, the exploration of which is a focal point in the present research. Related to that, the notion of learning (languages and other 'objects'), interestingly conceptualized as "*trajectories of ecological experience and repertoires of participation, gained in the process of adaptive dynamics*" (Atkinson et al. 2007: 172, italics in original), seems too bound up with the learning as participation metaphor to be relevant here. The predisposition towards learning as participation is stated programmatically in Atkinson (2002), but in Atkinson et al. (2007) the theoretical stance has been changed to include the acquisition metaphor as well. In

practice, however, this latter research, while not elaborating on either metaphor, studied only the situated nature of the learning of a certain construction/grammatical form<sup>95</sup> *have you ever*, without entering the discussion of how exactly learning happened as increased participation or how the grammatical form relates to other parts of the linguistic inventory. Furthermore, this study also claims language to be socially situated while also showing, albeit very briefly, the grammatical form under investigation to be transported by their focal student into other social situations, without discussing the implications of this for their views on language learning or language knowledge. To sum up, the sociocognitive approach, while sympathetic and promising, still seems to carry some terminological inconsistencies which render it impossible to implement for the present purposes.

Mondada and Pekarek-Doehler (2004), while not attempting to redirect the field of SLA, wrestle the problem of how to define the role of interaction in the process of learning. Combining the learning as participation metaphor from Lave and Wenger's (1991) situated learning theory with insights from SCT, they argue in favour of what they call a strong socio-interactionist position on L2 learning, willing to detach completely from the cognitive orientations of traditional SLA. Such a position entails the viewpoint that interaction is basic to and constitutive of all other kinds of human activity, thus making it the primordial locus for all aspects of social life, including (language) learning. Mondada and Pekarek-Doehler move on to show empirically how learning and the local anchoring in interaction are inseparable, as their data show how classroom learners (and their teachers) collaboratively and continually reconfigure the task-at-hand. Cognition is therefore viewed as socially situated in the local contingencies of everyday actions which renders the traditional SLA view of competence as isolated from the socialization processes impossible. As such, the strong CA-orientation, its roots in sociology, become clear, and the programmatic standpoints so far are taken well here. However, their approach and the present one seem to part ways when Mondada and Pekarek-Doehler (2004: 514) conclude on the basis of their investigations that their "analyses problematise assessing competence independently of social situations" and so "competence cannot be defined in purely individual terms". I fundamentally agree with the idea that the assessment of individual competence independently of aspects of sociality is problematic, but I would also maintain that it is possible to define competence as something retained by individuals as they navigate through the social world.

---

<sup>95</sup> Aktinson et al. use these terms seemingly interchangeably, which, unelaborated, is problematic from a UBL viewpoint, because of UBL's coercion of meaning and form in the 'construction'.

The same goes for language and language learners. The individual inventory, the collection of memories constituting the linguistic luggage of 'competence', is describable and a valid, if not desirable, object of investigation. We all have a human brain as accumulator but what goes in and out of it is not identical for all. Mondada and Pekarek-Doehler may not agree with this, but that does not make the enterprise impossible, as they claim. Arguably, they reach this conclusion because their conceptualisation of learning as participation, in extreme terms, views learning as locally managed and something that goes on as people co-achieve interactional completion. The data they forcefully use to back this view of learning do not, because they are not longitudinal, permit them to make claims about any portable skill that the 'learning as participation' might have resulted in. Therefore, even though I agree with Mondada and Pekarek-Doehler that their results make it impossible, or at least futile, to separate learning from the social world of the interactional here-and-now, their data do not support the dismissal of individual competence.

Lave and Wenger's (1991) situated learning theory itself does not, however, require an abolition of individual competence as research object. At the very core of the theory is not only the learning as participation metaphor, but also a 'learning as becoming' metaphor Wenger (1998). These metaphors are parallel; as a novice gradually moves from peripheral to fuller participation he also gradually becomes a fuller member of a given community of practice (Lave and Wenger 1991; Wenger 1998). Sfarid (1998), comparing the learning as participation metaphor and learning as acquisition metaphor, and warning against choosing either one at the cost of the other, described them as doing vs. getting / having. While language is something we arguably both do and have, for L2 research purposes we should heed Sfarid's warning; neither one works in terms of an exhaustive account of L2 learning. I try to show this empirically in ESK3. For SLA, however, as Larsen-Freeman (2004) notes in her taxonomy concerning learning metaphors, there has been a tendency that psycholinguistic SLA refers to learning as acquisition whereas sociolinguistic SLA (for want of a better term) investigates learning in terms of participation. This means that as a matter of preferred epistemologies, learning by implication becomes individualized in the former and socialized in the latter. The point I would like to make at this stage, however, is that neither learning metaphor necessarily excludes 'social context' or 'individual competence', respectively, from having relevance. This is done by researchers applying the respective terms. The reality of the mutually constitutive nature of language and context is not overruled by any metaphor of learning. Hall (e.g., 1993, 1995, 1997, 2004) in her Vygotskian approach to issues in L2 learning, while arguing in

favour of learning as becoming a competent member of a group, supports this reflexive nature of self and world; individual language behaviour and development, she argues, "originates in our socially constituted communicative practices" (Hall 1997: 302).

This problem is tackled by Hellermann (2007: 86) who notes that researchers working within the frameworks of sociocultural theory, situated practice, or CA are yet to establish "how to understand learning as individuals' change in participation in discursive practices over time". Hellermann's work (Hellermann 2006, 2007; Hellermann and Cole *forthc.*) fills a gap in the broad SLA field by systematically applying CA methods and situated learning (Lave and Wenger 1991) to longitudinal research in language learning. As such, Hellermann's work represents an important step towards investigating the cardinal questions for traditional SLA, namely development of individual competence over time. It must be stressed, however, that Hellermann operationalizes individual competence in a completely different manner than is traditionally done, as he views language learning as changes in interactional competence as displayed by the learners' participatory changes over time. He successfully shows how recurring classroom practices such as task openings and closings are fertile grounds for interactional competence development, as students demonstrably work out, in a collaborative manner, means to accomplish activities and reuse those means later in development to accomplish similar activities. An example of this (Hellermann 2007) is one student's appropriation of another student's linguistic means to assign interlocutors in a task-opening sequence. Initially Abby hears Frank openly ask for an interlocutor by publicly asking *who's talk to me?* in the classroom. In this interaction, a third student, Ana, responds and they engage in the task. At two later points, Abby is doing the interlocutor assignment, seemingly reusing the constructional format of Frank's original 'voice' – *I talk to you?, you talk to him* – as she assigns partnership to several students, including herself. A similar example comes from Hellermann and Cole (*forthc.*) where a learner is demonstrated to develop his interactional competence as his linguistic resources for achieving task-closing activities expand.

In some ways, Hellermann's research resembles the present research in that it captures the dynamic interplay between social context and L2 learning. However, the two research approaches also differ with respect to what exactly is under investigation. Hellermann's work investigates social practices that might be sites for language learning (Hellermann and Cole, *forthc.*), tracing similar social activities over time to explore the participants' change in participation in such practices over time.

The present research, on the other hand, traces linguistic patterns over time and tries to capture the linguistic interactions, usage events, in which those patterns are made relevant. The former approach offers insights into developing practices in the L2 classroom and the participants' organization of and orientation to such practices over time; the latter seeks to map out the contents of an emergent linguistic repertoire as it is put to use in classroom interaction. The two approaches, seen from the perspective of the present research, are therefore inter-compatible and complementary; this is implicitly underlined in Hellermann (2007) in which he states his sympathy towards a notion of grammar as emergent in nature. Combined, such approaches might, in the future, cast new light on the cardinal questions for SLA in that they are capable of investigating developing L2 competences in terms of a more holistic view on language and communication; a view which holds grammar to be both emergent and situated, both context-transcendent and locally contextualized.

Young and Miller (2004) travel down a parallel road to that of Mondada and Pekarek-Doehler, as well as Hellermann, viewing language learning in terms of changes in participation, and language and context as mutually constitutive along the lines proposed in the interactional competence framework. The situated learning theory (Lave and Wenger 1991), as Larsen-Freeman (2004) also notes, is increasingly widely applied in SLA research (e.g., Brouwer and Wagner 2004; Hellermann 2006, 2007; Hellermann and Cole *forthc.*). In Larsen-Freeman's (2004) taxonomy, it is made obvious that the participation metaphor is associated with research in which learning is conceptualised as increased membership, whereas the acquisition metaphor is associated with research in which learning concerns accumulation of a priori target rules. The present research fits into neither category, and therefore, in this respect, it makes a move toward combining the two learning metaphors; this is evident in the evolution of the problems investigated in the five research articles which increasingly attempt to explore the co-development of participatory patterns and linguistic units used in the classroom. In other words, the evolution of the present research, extremely data-driven as it is, shows that it is not built on any a priori construct of learning. Rather, the data have shown in an increasing number of instances, as reflected in the five research papers, the need for the application of both the learning as acquisition and learning as participation metaphors; the former blocks insights into socially informed nature of language learning and the latter is reluctant to speak of portability of linguistic patterns. Thus, the most fruitful, if not the only,

navigable road towards and understanding of the sociality of language learning and the ensuing portability of language abilities is cobbled with the stones of combining the two learning metaphors.

While both learning metaphors thus have a role to play, it would seem that research objectives to a great extent determine the choice of learning metaphor. While the participation metaphor has been ignored in traditional SLA, it is applied in L2 research which is not concerned with any form of accumulation of linguistic resources over time by individual learners (Young and Miller 2004; Hellermann and Cole *forthc.*). For the present purposes I could not argue that interaction is more important than cognition in a language learning situation; but I would have to maintain that neither is dispensable. Similarly I could not just describe and research learning as changes in participation patterns without also seriously considering the contribution made linguistically by the individual in social settings; in terms of SLA, this individual contribution is obviously more relevantly researched along the lines of a cognitive language inventory. Hence, it is possible to do away with a futile interaction vs. cognition discussion and enter the sphere of language use as the primordial setting for learning language as interaction and participation while remaining true to the idea of language and context as mutually constitutive.

In the learning as participation framework researchers have a tendency to dismiss the point that learners might also learn stuff of language as trivial to overall concerns of participation change. For example Hellermann and Cole (*forthc.*) speak of increased participation as their focal student develops in terms of disengaging from dyadic interactional tasks in the classroom, but, because they dismiss the idea of accumulation of linguistic material as part of their research in language learning, they fail to comment on the fact that in term of linguistic capacities he goes from saying nothing to saying *thank you* when disengaging, thereby displaying linguistic development. Similarly, Brouwer and Wagner (2004) explore interactional development in terms of establishing social relations and changes in participation without noticing that their focal participant displays changes in his deployment of German *haben*; over the course of three phone conversations with the same co-participant, he goes from saying *habscht Sie* to *haben Sie*, arguably, unless the idea of performance noise is posited, showing some form of linguistic development. This tendency is also evident in Young and Miller (2004) where the relation between language experience and participation is only implicitly mentioned in that minimal participation is coupled with the use of *yeah* only, whereas more increased participation is measured in terms of 'interactional competence' (Kramsch 1986;

Hall 1993), which can be summarized as language users' ability to locally co-construct a number of different routines needed in interactional practices and procedures (Hellermann 2007). Such practices and procedures, e.g., the turn-taking machinery of conversation (Sacks et al. 1974), are systematic, as language users we know them and obey them, we form habits along the lines of their prescriptions. If interactional competence is thought of as the ability to handle such procedures, and language learning is about acquiring them, linguistic units of some kind must lie at the heart of learning; after all, the turn-constructive units are the building blocks of turns (Sacks et al. 1974) and must by necessity carry linguistic information. It is therefore strange that Young and Miller simply wave off the fact that their informant deploys an increasing amount of linguistic items over time as unimportant. In other words, they acknowledge the presence in their data of evidence of some kind of linguistic development, but choose to ignore it because their interest lies elsewhere. They are entitled to see it that way, but stating up front that their research will contribute to our understanding of language learning they seem to be navigating contradictory waters in that they only deal with learning in terms of changing participation and interactional competence, but not individual linguistic resources. Young and Miller (2004), criticising the traditional view of L2 learning as development of mastery of linguistic form (also Mori 2004), tend to view language learning in purely social, behavioural terms, missing the concept of linguistic development entirely, which sets it apart from the research interests in the present research. From the perspective of my five research papers I cannot see how interactional competence could ever develop without being paralleled by some kind of individual linguistic development. These things are inextricably coupled.

This issue of contextualizing the nature of linguistic knowledge is approached, albeit from somewhat different angles, by Kasper (2004) and Larsen-Freeman (2004). Kasper discusses interlanguage in a narrow and broad definition, the former being exclusively about the development of linguistic competence, the latter concerning the development in L2 communicative competence. While communicative competence for SLA, as we saw earlier, definitely broadened the scope of competence per se in SLA, it did not result in a fundamentally different view of language knowledge than that prevalent at the time. The relationship between environment and interlanguage, on the other hand, is still largely uncharted territory, as Kasper (2004) reminded us. Even though the construct of interlanguage is problematic to the extent that it implies a distinct and stable endpoint, Kasper's message here is well-taken; the interplay between the environment of the classroom and my focal students is one of the core issues in ESK1, ESK3, and ESK4. Larsen-Freeman starts



her discussion from the point of difference between CA for SLA and traditional SLA; they have been kept apart, she argues, because the former is concerned with real-time establishing and co-constructing of identities and meaning and so on, whereas SLA has been concerned with a hypothesis-testing individual on the look-out for portable language structures. Arguably though, there is no contradiction, really. It is more of a continuum; in order to fully grasp SLA, both ends of this continuum need attending to. Cognition is locally applied but may also transcend context. The same thing applies to linguistic constructions. Nothing is only local or global; everything is potentially both; in languages and elsewhere.

This line of argumentation seems to culminate in the idea that the psychological reality of linguistic units posited in SLA contrasts with CA's "allow[ing] units to emerge from the data" (Larsen-Freeman 2004: 605). While it is unclear what exactly is meant by this statement, it might mark a contrast, as Larsen-Freeman explains, between psycholinguistic SLA and CA for SLA. But, keeping tongue-in-cheek, this does not fundamentally mark a contrast between CA and cognitive approaches to the study of language. It might mark a contrast between CA and Formalism because the units they each find attractive for linguistic analyses are fundamentally in opposition; as Markee and Kasper (2004) point out, the former deals with clauses (and, I might add, utterances) as relevant units, whereas the latter is based on the sentence as relevant unit. So the difference lies in assumptions about language, not assumptions about interaction or data (though such empirical issues are all interrelated). The interesting thing is, as has been implied earlier, that this is a matter of being sufficiently data-driven. This is a matter of approaching data from a more emic perspective than usually done in SLA and it is a matter of viewing the unit of learning in a different way than is usually done. The contrast is there due to long-standing epistemologies in each field, which Larsen-Freeman on a more abstract level also concedes. The room is there for combining linguistic theory, learning theory, and micro-analytic procedures of social interaction. The door into that room is UBL with its empirically based linguistic apparatus and sincere interest in what people do with language – and how they learn it while doing so. Such questions cannot be answered without minimally responding to issues of linguistic structure, learning metaphors, and interactional analysis. Larsen-Freeman (2004: 607) herself encourages the field to apply both learning as participation and learning as acquisition metaphors, and to "look to the union of use and acquisition", thus eventually echoing Firth and Wagner (1998), even though she also seems to maintain that there are profound differences between psycholinguistic SLA and "CA for SLA", primarily because of CA's social

view on language and learning and SLA's traditional mentalistic views on these matters. In a 2007 response to the same debate, Larsen-Freeman, in what seems to be a small and natural step for her but which might be a great leap forward for SLA, finally seems ready to let go of past caveats; she now means to assist SLA in taking a step further as she states her predisposition towards an appreciation of where approaches meet and interact to overcome the dualistic nature of many aspects of the field; the same dualistic nature of the field, of course, under investigation here, and explored to varying degree in ESK1, ESK2, ESK3, and ESK4.

#### 8.4 Overcoming the dualisms

Bringing this discussion to an end and in order to motivate it in light of the present research, the position taken here on the dualistic nature of many aspects of SLA is that the line of thinking along such lines of shooting the world in black and white should be discontinued. It makes no sense to speak of mutual exclusivity of knowledge and use, acquisition and use, generality and locality of linguistic items, acquisition and participation, etic and emic perspective on L2 data, and sociality and individuality. It has been shown that the IIO-model and its accompanying "negotiation for meaning" view of interaction, is based on false pre-conceptions of what it entails to learn a language. There simply is no empirical validity to the claim that interaction as it happens between native speakers and non-native speakers propels acquisition, the point of uptake of linguistic structures unestablishable, nor is it empirically sound to posit that language learning tasks, which enhance a particular learning objective, are better learning contexts than regular, free conversations (Kasper 2004). The acquisition-use dichotomy may thus be dismissed as empirically invalid, which invites for SLA investigations based on a conflation of the two (see Larsen-Freeman 2007 for a compatible view).

The present research explores L2 learning from the perspective of not dichotomising acquisition and use; it does not speak of 'knowing' as a finite state, but as a transitory ability, thus finding a distant forefather in Hymes (1972). With respect to language specifically, then, knowledge is experiential, an inventory of past memories of linguistic goings-on. Biographical language learning in this orientation builds on a person's interactions in the language. The development of an individual linguistic inventory is inextricably linked with that individual's interactional history; as Ohta (2001) notes, these two things have been kept apart in traditional SLA. Ohta herself validates this statement empirically by showing how students' private speech in classroom interaction evolves

into 'real' participation, public speech, as it were, in interaction over time. In my research papers, ESK1, ESK2, ESK3, and ESK4, this relationship between interactional requirements and linguistic inventory possibilities is shown by highlighting (ESK1) the mutually constitutive nature of MWEs and interactional routines, (ESK2) the emergence of utterance schemas in interaction, (ESK3) the use of a given MWE in multiple interactional settings in a participant-relevant way, and (ESK4) the situated development of one student's deployment of negation patterns; all investigations which point to the indivisibility of interaction and learning by supporting Firth and Wagner's (1998) tenet that it is impossible to pinpoint exactly where one ends and the other begins. Correspondingly, the approach taken here does not give prominence to sociality over individuality. Given the hunt for the stuff of individual linguistic inventories, it could be argued that individuality is emphasised at the cost of sociality. Granted, the present research is not particularly interested in social dynamics of interaction; rather, it is focused on finding a relevant unit for investigating the individual portability of language from one context to another (Larsen-Freeman 2004). These contexts, however, are explicitly social and the development of the individual inventories has been shown to be directly dependent on the activities they provide. Such a starting point is much more fertile ground than the IIO-model which states up front that it is not interested in use (i.e. interaction) but in acquisition. In all fairness, they should have left out the 'interaction' part of their model and called it 'the input-model', or 'the meaning-negotiation model' or something to that effect. It has nothing to do with interaction.

The present research has resulted in an understanding of the linguistic knowledge involved in L2 learning as an emergent repertoire of routines and utterance schemas. This implies that locality and generality of deployment both need to be taken into account in analyses and investigations of linguistic development. It also means that it is necessary to embrace an emic perspective on the data. The locally contextualized nature of the view of language learning propagated here is most profitably investigated in terms of interactional participant-relevance. It was shown, e.g., in ESK3, that issues of sequentiality and interlocutors' orientation fundamentally impact the way utterances are understood and ongoing interaction is projected. It has also been shown that some aspects of the linguistic inventories of my two focal students are transported through time to be deployed at later times. Most prominently, this feature is the prerogative of utterance schemas, i.e. patterns that are part stable, part productive, the latter part conceptualised as schematically sanctioned. Such an understanding of the linguistic inventory, building on usage-based linguistics, seems to sit well with

CA, not only because linguistic units of UBL may be parallel with CA's turn-constructive units, but because of the demonstrably interactionally situated nature of much language. Even though CA and UBL are epistemologically different, the former interested in social order and social action, the latter interested in linguistic knowledge, this intuition is supported in the writings of some socially oriented SLA researchers; e.g., Kasper (2004), who speaks of repertoires of her focal students in a manner which sounds like vintage UBL, and Hall (2004: 610) who, referring to UBL researchers Hopper and Tomasello, speaks of language knowledge as a "constantly evolving set of recurring regularities". Formulations such as Hall's give hope for the emergence of a new performance-driven SLA eclecticism where researchers do not blindly follow ahistorical convictions but instead set new empirical standards for their investigations. Data vigour should inspire the formation of new categories for analyses, not theoretical hegemony. The field needs a wider frame of reference (e.g., Larsen-Freeman 2002, 2007; Watson-Gegeo 2004) so as to incorporate into its main machinery new ways of practicing eclecticism. So, *e pluribus unum*? No, plurality will and should prevail.



## Chapter 9

### Conclusions and implications.

What was from the outset framed around a wish to explore the role of 'formulaic language' in L2 learning from the perspective of functional-cognitive linguistics quickly became a wish for exploring the fruitfulness of applying the framework of Usage-Based Linguistics (UBL) to general investigations into L2 learning. The question of whether L2 learners start out from formulas and gradually start analysing them to use the individual constituents in other linguistic structures, or if they start from the learning of combinatorial rules with practice ensuring the entrenchment of certain formulas over time, seemed fundamentally misguided and flawed. Rather, formulas and more general patterns were found to co-exist at all points in development, at least as far as the data and focal students investigated here are concerned.

What has been invoked, therefore, as a guiding principle in the research carried out here is the UBL path of learning from formulas via partially schematised patterns to fully abstract constructions. This modus operandi proved extremely useful and empirically valid from the beginning. In the first research paper, in which the UBL framework was applied, EC, we found positive evidence for the UBL learning trajectory, with the system emerging in acquisition characterized as the gradual abstraction of regularities that link expressions as constructions. *Do-negation* learning was found to be initially heavily reliant on one specific instantiation of the pattern, *I don't know*, with productivity gradually increasing as the underlying knowledge seemed to become increasingly abstract, as reflected in type and token frequencies. *I don't know* was also found to be stable throughout development suggesting its entrenchment as a MWE. The findings suggested that L2 learning is indeed item-based, that expression entrenchment is dependent on token frequency (as in the case of *I don't know*) and that more abstract pattern and construction learning is dependent on type frequency, as expected.

The item-based nature of pattern development as predicted by the UBL path of learning was then further investigated and supported empirically in subsequent papers, especially ESK1, ESK3, and ESK4. The item-based development could be demonstrated for the various patterns under investigation:

- The general *do-negation* pattern emerged from *I don't know* to become increasingly varied and productive. This was valid for both my focal students.
- Valerio's learner pattern *you no verb* was also item-based as it was found to be initially dependent on a local high recurrence of *you no write*. This pattern was also argued to momentarily take the lead, as it were, in the competition against the target-language variety.
- This target-language variety, in turn, was dependent on two things: 1) a previously recurring pattern as *you no useh* may have evolved into *you don't useh*; and 2) the locally recurring *you don't know* – of course, the already highly frequent *I don't know* a possible psycholinguistic influence.
- Valerio's *aux-do*-pattern was also found to be item-based, initially heavily dependent on the MWE *what do you say (for)*, and gradually spreading to work with a range of other main verbs
- *Can*-usage by Carlos was a more tricky matter, but could be described as being item-based in a highly locally contextualized fashion, springing from a briefly recurring use of *I can write*.

In turn, and chronologically, the research papers each brought about insights which spawned new research questions. ESK1 investigated the item-based nature of Carlos's *can*-patterns, but failed to yield a result as stream-lined as had been the case in EC. In that research, Carlos's learning of the *do-negation*-pattern had been shown to be in complete alignment with the UBL path of learning. In ESK1, it transpired that formulas, or multi-word expressions (MWEs), should be seen as interactionally and locally contextualized. They were found to be transitory in nature; i.e. their deployment over time was seen to be occasioned by specific usage events. Such events, it was argued, must be recurrent in order for the MWEs, at least the ones identified for Carlos's *can*-pattern development, to be retained by the learner over time. It was further observed that productivity enhancement is partially concrete, based on utterance schema development, and traceable to previous experience. The traceability of linguistic patterns in development to previous experience was depicted in terms of the emergentist stepping-stone metaphor, which captures the experiential linguistic development as it happens on a pattern-by-pattern basis as the learner, Carlos, constructs his L2 inventory.

It was suggested that differences in findings between EC and ESK1 were a matter of the linguistic pattern under investigation. The *negation* construction, it was speculated, lends itself more easily to schematic abstraction than the *can*-pattern. MWEs in use for the *can*-pattern also seemed more situationally dependent than the MWE *I don't know*; they were found to be less generally applicable in use. It was therefore argued that analyses of emergent linguistic inventories should also take the most concrete starting point possible; matters of schematic abstraction and generality of use should be empirically substantiated rather than assumed a priori. It therefore lies at the heart of the findings in ESK1 that development is not only item-based, but also very much usage-based. It was displayed how patterns wax and wane (Thelen and Bates 2003; Larsen-Freeman 2006) in response to changing environmental and interactional factors. It follows from these insights that ontogenetic language development is inextricably coupled with language use. A full theory or model of SLA, it was therefore argued, must incorporate room for studying these local contexts in a more detailed manner to investigate in depth the interplay between local interactional contingencies and portable linguistic experience. The results in ESK1, then, yielded the research questions to be tackled in the subsequent research papers. In more concrete terms, the question that informed ESK2 lay in delineating the characteristics of those aspects of language knowledge that are situated and transitory and those that are durable and portable. ESK1 indicated that MWEs may be generally transitory and locally contextualized whereas more schematic language knowledge may be less susceptible to environmental changes and thus more sturdy in its portability.

Inspired partly by ESK1 and a study in child language acquisition by Lieven et al. (2003), ESK2 attempted to delineate issues of locally contextualized routines and more general creativity. ESK2 understood the object of research in longitudinal L2 studies as a hybrid between locally applied usage patterns and application of the same and related usage patterns over time, and represents the germ of the idea of viewing L2 learning in terms of an empirically grounded, emergent grammar, consisting of units of spontaneously occurring language use. This resulted in a conceptualisation of emergent creativity as building on recycled linguistic matter in the form of MWEs and utterance schemas, i.e., patterns that are more or less lexically specific.

ESK2 also empirically substantiated L2 development to be too complex and non-linear, the various utterances perhaps not learnt by way of formal causality, to be exhaustively captured by strict terms of syntactic operations which in the study by Lieven et al. had proven to be a very fruitful approach



to early L1 development. It was shown that L2 development is not that easily captured, that the linguistic inventory (in a L2) seems to be a structured set of utterance schemas. Development was thus described and analysed as the emergence of new utterance schemas and the combination of such schemas, in an increasing number of ways with an increasing number of schematically sanctioned lexical options and intra-turn schematic operations.

In ESK2 it was also shown how combinability is about putting together chunks rather than lexical items as separate islands. In other words, the lexical items employed are dependent on the patterns known to the language user – and the patterns seem to have been learnt in lexically specific environments as item-based. This insight, it was argued, equals an empirically substantiated justification of ruling out syntactic combinability as the stuff of learning; instead, item-based utterance schemas were posited as the main linguistic material to learn. Pawley and Syder's puzzle of native-like selection, it was argued, could be solved by leaving behind once and for all the dualism inherent in the lexicon-syntax division and focus, teachers, learners, users, researchers alike, on a description and analysis of linguistic inventories as item-based as proposed in usage-based linguistics.

ESK2, then, paved the way towards an understanding of the object of research in longitudinal L2 studies as a hybrid between locally applied usage patterns and application of the same and related usage patterns over time; i.e., an empirically grounded, emergent grammar, consisting of units of spontaneously occurring language use. As such, it stands as a corner-stone in the present research. However, the subsequent research papers were actually to a greater extent fuelled by several questions from ESK1 which had remained unapproached in ESK2, namely questions pertaining to the interdependent nature of linguistic interaction and language learning.

ESK3 showed how the use of a MWE, *what do you say*, was initially situated in a recurring environment but later expanded to be used in other environments as well. It was argued that the identical deployment of the utterance in comparable sequences over time, an initial routinisation of the MWE, was an example of Valerio having been socialized into the language classroom practice of inviting for help (Brouwer 2003). This, in turn, was seen to hang together with the learning as participation metaphor, whereas Valerio's ability to use the MWE in a new context was argued to be more profitably thought of in terms of the learning as acquisition metaphor. Based on his

activities in the social world of the classroom practices, then, he improved his productivity. As such, it was demonstrated how a full account of L2 learning needs to take into consideration both participation in social interaction and psycholinguistic notions of cognitive portability, linguistic resources, in terms of acquisition.

Furthermore, it was demonstrated how these learning issues fit nicely into a rethinking of the performance-competence distinction along a time dimension of local performance enhancement (routinisation) and general productivity enhancement which transcends the moment. This introduced an elaboration of Larsen-Freeman's (2004) conceptualization of learning as that which is carried across contextual boundaries. *What do you say* was found to be carried across contextual boundaries, time-wise, but only so to a certain extent content-wise; i.e., the expression is locally, interactionally contingent, it does not become relevant in any old conversational situation. The utterance schema that emerges from the MWE, namely the more general *do-schema*, while thus a sediment of those interactional contingencies, is much more generally applicable and not, in terms of use, dependent on a narrowly defined conversational setting. The utterance schema, therefore, is carried across both content-defined and time-defined contextual boundaries. This, in essence, marks the difference, fluid as it is, between 'performance enhancement' and 'productivity enhancement' as I operationalised them in ESK3. Looking back, this distinction between a time-defined and a content-defined notion of pan-contextual portability also seems to apply to the findings in ESK1, where the situated MWEs all seemed to be coupled with certain interactional requirements; i.e., they displayed time-wise portability did not seem to be portable into new environments.

ESK4 further supported the interrelationship between linguistic development and interactional requirements. It was shown that Valerio, in usage events requiring him to assist his fellow classmates in getting a task right, used the linguistic resources readily available to him; namely, a non-native-like, lexically specific pattern *you no write*. It was argued that in terms of the UBL framework and the importance it ascribes to issues of type and token frequencies, these usage events which prompted the locally heavy use of *you no write*, may have laid the foundation for what in Valerio's ontogenesis was to become a seemingly statistical feature of his linguistic inventory, namely the co-occurrence in negation patterns of *you* and *no* at the cost of a more native-like *do-negation* pattern. ESK4, then, yet again showed the futility of keeping interaction and learning apart.

The developmental tendencies found in the data for ESK4, *do-negation* patterns for both my focal students, were found to support some of the findings, and refute others, from research in developmental sequences. More importantly, ESK4 questioned the fundamental starting point for defining those sequences, as the data supported the finding from ESK2 that learning L2 syntax is not a matter of context-independent rule-learning across linguistic patterns; rather, it is a matter of construction-dimensional exemplar-deduced tendencies that may or may not become schematized as abstract linguistic knowledge in ontogenesis. It is possible to track pattern development in great detail from the concrete item-based starting point of the patterns to the possible abstraction of regularities that link these patterns as schemas. Such possible abstraction, however, should not be the default starting point for longitudinal L2 learning studies, because, as ESK1 showed, not all patterns lend themselves easily to abstraction.

On a more epistemological note, the application of UBL and its insistence on real usage in real usage events as basis for research has resulted in an empirical substantiation of the futility of keeping 'learning' and 'use' apart in SLA studies. It simply is, as Firth and Wagner (1998) told us, impossible to tell where one ends and the other begins. L2 learning is not as simple as being a matter of restructuring cognitive machinery against the influence of attended input in modified interaction. The model assuming these basics for L2 learners and L2 researchers like, the Interaction/Input/Output-model, has been shown to be inadequate for dealing with the complexities of learners constructing their L2 inventories. Rather, what the data suggest is that linguistic development and interactional requirements are interwoven at a fundamental level; what learners learn is not the outcome of what they do, it is something that emerges in the flux of doing. Learning and doing are mutually reflexive; they happen simultaneously. Some cases in point are the emergence of *you don't verb* in ESK4, the gradual spreading of the *can*-pattern in ESK1, as well as some of the emergent patterns documented in ESK2 (*in + location* and *it's more X*).

Furthermore, ESK3 and ESK4 showed the need for an elaborate investigative framework to tackle the interactional phenomena encountered in order to capture the essence of the interplay between interaction and learning. For this purpose, the notion of usage event as it is found in UBL did not seem sufficient. Therefore, I invoked Conversation Analysis (CA), or at least some micro-analytic tools inspired by CA, to account for local interactional contingencies found in the data. This

resulted in an eclectic approach to the study of developmental issues in L2 learning, suggesting the need to abolish or rethink certain dualisms; apart from the learning-use dichotomy, these included the performance-competence distinction and the dichotomous relationship between the two learning metaphors, 'participation' and 'acquisition'. It was argued that developmental trajectories found in the data should be analysed and described with recourse to both metaphors, and it was argued, and put into further perspective in chapter 8 on SLA, that an eclectic approach, which attempts to bridge the gaps represented by existing dualisms, is more apt at accounting for L2 learning than one which chooses either side of the dualism fence.

Summing up, it has been argued that UBL is a fruitful framework for exploring developmental issues in L2 learning. The five research papers have, each from different perspectives and with different research objectives, substantiated the UBL claim that language learning is item- and usage-based. The usage-based perspective, increasingly pivotal in the chronology of the five research paper, has called for a non-distinction between interaction and learning. Rather, these have been argued to be fundamentally coupled and happen simultaneously. Analytic tools from CA were found to be immensely useful to analyse the learning trajectories which led to these conclusions. The item-based nature of the projected trajectory of language learning has implied a view of 'formulaic language' which is different from that found in traditional psycholinguistic research (e.g., Wray 2002; Schmitt and Carter 2004) which has been argued to subscribe to a view of language knowledge as essentially compartmentalized, entailing a definition of formulaicity as something not generated by syntax. UBL, on the other hand, does not distinguish the phenomenon as essentially different from the rest of the linguistic inventory. Language knowledge is fundamentally holistic, all patterns of language uniformly represented in the inventory (Croft and Cruse 2004). This view of language knowledge has been empirically substantiated in L1 acquisition research; Tomasello (2003) put it nicely when he said for children constructing their first language that they must learn two faces of grammar: smaller elements and larger patterns. Now it seems that it is time for the field of SLA to apply this insight and investigate its validity in ever more detail for research on adults accumulating their linguistic resources as they construct their L2 inventories. SLA must rid itself of the compartmentalized view of language; i.e., the idea that lexis and grammar are to be kept apart, in research as well as in teaching. L2 learners simply do not learn the two in a manner that justifies keeping them apart; they are intertwined to the brink of being inseparable, and learners acquire them together, not each in its own paradigmatic vacuum.

## Complete References.

- Achard, M. 1997. Cognitive Grammar and SLA investigation. *Journal of Intensive English Studies* 2: 157-176.
- Achard, M. 2006. Usage-based semantics: Meaning and distribution of three French 'breaking' verbs. *Paper presented at Collocations and Idioms: The First Nordic Conference on Syntactic Freezes, Joensuu 19-20/5-06.*
- Achard, M. 2007. Usage-based semantics: Meaning and distribution of three French 'breaking' verbs in Nenonen, M and S. Niemi (eds.). 2007. *Collocations and Idioms 1: Papers from the First Nordic Conference on Syntactic Freezes, Joensuu, May 19-20, 2007.* Studies in Languages, University of Joensuu, Vol. 41. Joensuu: Joensuu University Press.
- Achiba, M. 2003. *Learning to request in a second language: A study of child interlanguage pragmatics.* Clevedon: Multilingual Matters.
- Adolphs, S. and V. Durow. 2004. Social-cultural integration and the development of formulaic sequences. In: Schmitt, N.. (ed.) 2004. *Formulaic sequences.* Amsterdam: John Benjamins.
- Albrechtsen, D., K. Haastrup, and B. Henriksen. 2008. *Vocabulary and Writing in a First and Second Language. Process and Development.* Houndmills: Palgrave Macmillan.
- Alexander, R. J. 1985. Phraseological and pragmatic deficits in advanced learners of English: Problems of vocabulary learning? *Die Neueren Sprachen* 84/6: 613-21.
- Appel, R. 1996. The lexicon in second language acquisition. In: Jardens, P. and J. Lalleman (eds.) *Investigating Second Language Acquisition* (Berlin: Mouton de Gruyter).
- Atkinson, D. 2002. Toward a sociocognitive approach to second language acquisition. *The Modern Language Journal* 86/4: 525-545.
- Atkinson, D. 2007. Toward a sociocognitive approach to second language acquisition. Colloquium talk presented at AAAL, Costa Mesa, California, 21-24 April 2007.
- Atkinson, D., E. Churchill, H. Okada, and T. Nishino. 2007. Alignment and interaction in a sociocognitive approach to second language acquisition. *The Modern Language Journal* 91, 2: 169-188
- Austin, J. L. 1962 *How to do things with words.* Oxford: Clarendon
- Bahns, J. and M. Eldaw. 1993. Should we teach EFL students collocations? *System* 21/ 1: 101-114
- Bardovi-Harlig, K. 2002. A new starting point? Investigating formulaic use and input in future expression. *Studies in Second Language Acquisition* 24/2: 189-198.

- Bardovi-Harlig, K. 2006. On the role of formulas in the acquisition of L2 pragmatics. In: Bardovi-Harlig, K., C. Félix-Brasdefer, and A. Omar (eds.): *Pragmatics and Language Learning, Vol. 11* Honolulu, HI: National Foreign Language Resource Center, University of Hawai'i at Manoa.
- Barlow, M. and S. Kemmer. (eds.). 2000. *Usage-Based Models of Language*. Stanford: CSLI.
- Barlow, M. and S. Kemmer. 2000. Introduction: A usage-based conception of language. In: Barlow, M. and S. Kemmer. (eds.). 2000. *Usage-Based Models of Language*. Stanford: CSLI.
- Bates, E. and B. MacWhinney. 1982. Functionalist approaches to grammar. In: Wanner, E. and L.R. Gleitman (eds.): *Language Acquisition – the State of the Art*. Cambridge: Cambridge University Press.
- Beheydt, L. 1987. The semantization of vocabulary in foreign language learning. *System* 15/1: 55-67.
- Beretta, A. 1991. Theory construction in SLA: complementarity and opposition. *Studies in Second Language Acquisition* 13/4: 493-511.
- Bialystok, E. and K. Hakuta 1994. *In other words – The science and psychology of second language acquisition*. New York, NY: BasicBooks.
- Biskup, D, 1992. L1 Influence on learners' renderings of English collocations: A Polish/German empirical study. Arnaud, P. J. L. and H. Béjoint (eds.): *Vocabulary and Applied Linguistics*. London: MacMillan.
- Bley-Vroman, R. 1983. The comparative fallacy in interlanguage studies: The case of systematicity. *Language Learning* 33/1: 1-17.
- Block, D. 1996. Not so fast. Some thoughts on theory culling, relativism, accepted findings, and the heart and soul of SLA. *Applied Linguistics* 17/1: 63-83.
- Block, D. 2003. *The social turn in SLA*. Edinburgh: Edinburgh University Press.
- Bogaards, P. 1996. Lexicon and grammar in second language learning. In: Jordens, P. and J. Lalleman (eds.): *Investigating Second Language Acquisition*. Berlin: Mouton de Gruyter.
- Bogaards, P. 2001. Lexical units and the learning of foreign language vocabulary. *Studies in Second Language Acquisition* 23/3: 321-344.
- Bohn, O.-S. 1986. Formulas, frame structures, and stereotypes in early syntactic development: Some new evidence from L2 acquisition. *Linguistics* 24: 185-202.
- Bolander, M. 1989. Prefabs, patterns and rules in interaction. In: Hyltenstam, K. and L. K. Obler (eds.): *Bilingualism Across the Life Span*. Cambridge: Cambridge University Press.
- Bolinger, D. 1979. Meaning and memory. In: Haydu, G. G. (ed.): *Experience Forms. Their Cultural and Individual Place and Function*. The Hague: Mouton.

- Bonk, W.J. 2000. Testing ESL learners' knowledge of collocations. *ERIC Documents 442 309*
- Brillianceau, D. 2005. Spontaneous conversation. A window into language learners' autonomy. *Focus on Basics 8a*: 22-25. URL: [http://www.ncsall.net/fileadmin/resources/fob/2005/fob\\_8a.pdf](http://www.ncsall.net/fileadmin/resources/fob/2005/fob_8a.pdf)
- Brouwer, C. E. and J. Wagner, 2004: Developmental issues in second language conversation. *Journal of Applied Linguistics 1/1*: 29-47.
- Brouwer, C. E. 2003. Word searches in NNS-NS interaction: Opportunities for language learning? *The Modern Language Journal 87/ 4*: 534-545.
- Brouwer, C. E. 2004. Doing pronunciation: A sequence of novice-expert interaction. Gardner, R. and J. Wagner (eds.): *Second language talk*. London and New York: Continuum.
- Brown, G. 1996. Language learning, competence and performance. Brown, G., K. Malmkjær, and J. Williams (eds.): *Performance and Competence in Second Language Acquisition*. Cambridge: Cambridge University Press.
- Butler, C. S. 2003. Multi-word sequences and their relevance for recent models of Functional grammar. *Functions of Language 10/2*: 179–208.
- Bybee, J. and P. Hopper. 2001. Introduction to frequency and the emergence of linguistic structure. In: Bybee, J. and P. Hopper (eds.): *Frequency and the Emergence of Linguistic Structure*. Amsterdam: John Benjamins.
- Cadierno, T. and L. Ruiz. 2006. Motion events in Spanish L2 acquisition. *Annual Review of Cognitive Linguistics 4*: 183–216.
- Cadierno, T. 2004. Expressing motion events in a second language: A cognitive typological approach. Achard, M. and S. Niemeier (eds.): *Cognitive linguistics, second language acquisition and foreign language pedagogy*. Berlin: Mouton de Gruyter.
- Canale, M. 1983a. From communicative competence to communicative pedagogy. Richards, J. C. and R. W. Schmidt (eds.): *Language and Communication*. Harlow: Longman.
- Canale, M. 1983b. On some dimensions of language proficiency. Oller, J. W. (ed.): *Current Issues in Language Testing Research*. Rowley Mass.: Newbury House.
- Canale, M. and M. Swain 1980. Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics 1/1*: 1-47.
- Carroll, S. E. 2002. Induction in a modular learner. *Second Language Research 18/2*: 224-249.
- Carroll, S. E. 2007. Autonomous induction theory. In: VanPatten, B. and J. Williams (eds.): *Theories in Second Language Acquisition. An introduction*. Mahwah, NJ.: Lawrence Erlbaum.
- Carter, R. 1987. *Vocabulary – applied linguistic perspectives* London: Allen and Unwin.

- Carter, R. and M. McCarthy (eds.). 1988. *Vocabulary and language teaching*. Harlow: Longman.
- Celce-Murcia, M., Z. Dörnyei, and S. Thurrell 1995. Communicative competence: A pedagogically motivated model with content specifications. *Issues in Applied Linguistics* 6/2: 5-35.
- Channell, J. 1988. Psycholinguistic considerations in the study of L2 vocabulary acquisition. Carter, R. and M. McCarthy (eds.): *Vocabulary and language Teaching* Harlow: Longman.
- Chomsky, N. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. 1980. *Rules and representations*. New York: Columbia University Press.
- Chomsky, N. 1986 *Knowledge of language – Its nature, origin, and use*. Westport, CT: Praeger.
- Chomsky, N. 1988 *Language and problems of knowledge*. Cambridge, MA: MIT Press.
- Chomsky, N. 2000. *New horizons in the study of language and mind*. Oxford: Oxford University Press.
- Clahsen, H. 1985. Profiling second language development: A procedure for assessing L2 proficiency. In: Hyltenstam, K. and M. Pienemann (eds.): *Modelling and Assessing Second Language Acquisition* Clevedon: Multilingual Matters.
- Clahsen, H. and P. Muysken 1986. The availability of universal grammar to adult and child learners – a study of the acquisition of German word order. *Second Language Research* 2/2: 93-119.
- Cook, V. 1993. *Linguistics and second language acquisition*. London: MacMillan.
- Cook, V.J. 1985. Chomsky's universal grammar and second language learning. *Applied Linguistics* 6/1: 1-18.
- Cook, V.J. and M. Newson. 1998. *Chomsky's universal grammar*. Oxford: Blackwell, 2<sup>nd</sup> ed.
- Corder, S. P. 1967. The significance of learners' errors. *International Review of Applied Linguistics in Language Teaching* 5/4: 161-170
- Cowie, A. P. 1988. Stable and creative aspects of vocabulary use. In: Carter, R. and M. McCarthy (eds.): *Vocabulary and Language Teaching*. Harlow: Longman.
- Cowie, A. P. 1992. Multiword lexical units and communicative language teaching. In: Arnaud, P. J. L and H. Béjoint (eds.): *Vocabulary and Applied Linguistics*. London: MacMillan.
- Croft, W. 1995. Autonomy and Functional linguistics. *Language* 71: 490-532.
- Croft, W. 2001. *Radical Construction Grammar*. Oxford: Oxford University Press.
- Croft, W. and D. A. Cruse 2004. *Cognitive Linguistics* Cambridge: Cambridge University Press.



- Dabrowska, E. 2000. From formula to schema: The acquisition of English questions. *Cognitive Linguistics* 11/1-2: 83-102.
- Dabrowska, E. and E. Lieven. 2005. Towards a lexically specific grammar of children's question constructions. *Cognitive Linguistics* 16/3: 437-474.
- De Bot, K., W. Lowie and M. Verspoor. 2007. A dynamic systems theory approach to second language acquisition. *Bilingualism: Language and Cognition* 10/1: 7-21.
- DeCarrico, J. S. 1998. Syntax, lexis and discourse: Issues in redefining the boundaries. In: Haastrup, K. and Å. Viberg (eds.) *Perspectives on Lexical Acquisition in a Second Language* Lund: Lund University Press
- Dik, S.C. 1989. *The Theory of Functional Grammar. Part 1: The Structure of the Clause* Dordrecht: Foris Publications
- Dörnyei, Z., V. Durow, and K. Zahran. 2004. Individual differences and their effects on formulaic sequence acquisition. In: Schmitt, N.. (ed.) 2004. *Formulaic sequences*. Amsterdam: John Benjamins
- Doughty, C. 2001. Cognitive underpinnings of focus on form. In: Robinson, P. (ed.): *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press
- Doughty, C. and M. Long 2003. Introduction. In: Doughty, C. and M. Long (eds.): *Handbook of Second Language Acquisition*. Malden, MA: Blackwell
- Du Bois, J. W. 2003. Discourse and grammar. In: Tomasello, M. (ed.): *The New Psychology of Language, volume 2*. Mahwah, New Jersey: Lawrence Erlbaum
- Ellis, N. C. 1996. Sequencing in SLA – phonological memory, chunking, and points of order. *Studies in Second Language Acquisition* 18: 91-126
- Ellis, N. C. 1997. Vocabulary acquisition; word structure, collocation, word-class, and meaning In: Schmitt, N. and M. McCarthy (eds.): *Vocabulary: Description, acquisition, and pedagogy* Cambridge: Cambridge University Press.
- Ellis, N. C. 1998. Emergentism, connectionism, and language learning. *Language Learning* 48/4: 631-664.
- Ellis, N. C. 2002. Frequency effects in language processing – a review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition* 24/2: 143-188.
- Ellis, N. C. 2004. The processes of second language acquisition. In: VanPatten, B., J. Williams, S. Rott and M. Overstreet (eds.): *Form-Meaning Connections in Second Language Acquisition* Mahwah, NJ: Lawrence Erlbaum.

- Ellis, N. C. 2007. The associative-cognitive CREED. In: VanPatten, B. and J. Williams (eds.): *Theories in Second Language Acquisition. An introduction*. Mahwah, NJ.: Lawrence Erlbaum.
- Ellis, N. C. and D. Larsen-Freeman. 2006. Language emergence: Implications for applied linguistics – Introduction to the Special Issue. *Applied Linguistics* 27/4: 558-589.
- Ellis, R. 1984. Can syntax be taught? A study of the effects of formal instruction on the acquisition of WH Questions by Children. *Applied Linguistics* 5/2: 138-155.
- Ellis, R. 1985/1996. *Understanding Second language Acquisition*. Oxford: Oxford University Press.
- Ellis, R. 1990. A response to Gregg. *Applied Linguistics* 11/ 4: 384-391.
- Ellis, R. 1992. Learning to communicate in the classroom: A study of two learners' requests. *Studies in Second Language Acquisition* 14: 1-23.
- Ellis, R. 1994. *The Study of Second language Acquisition*. Oxford: Oxford University Press .
- Ellis, R., Y. Tanaka, and A. Yamazaki. 1994. Classroom interaction, comprehension, and the acquisition of L2 word meanings. *Language Learning* 44: 449-491.
- Eskildsen, S. W. and J. Wagner. 2007. What counts as evidence in longitudinal language learning.? Paper presented at AAAL, Costa Mesa, California, 21-24 April 2007.
- Færch, C., K. Haastrup and R. Phillipson 1984. *Learner Language and Language Learning*. Copenhagen: Gyldendal.
- Ferguson, C. 1975. Toward a characterisation of English foreigner talk. *Anthropological Linguistics* 17/1: 1-14.
- Fillmore, C. J., P. Kay, and M. C. O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: The Case of Let Alone. *Language* 64/3: 501-538.
- Firth, A. and J. Wagner. 1997. On discourse, communication, and (some) fundamental concepts in SLA research. *The Modern Language Journal* 81/3: 285-300.
- Firth, A. and J. Wagner. 1998. SLA Property: No trespassing! *The Modern Language Journal* 82/1: 91-94.
- Firth, A. and J. Wagner. 2007. S/FL Learning as a social accomplishment: Elaborations on a 'reconceptualized' SLA. *Modern Language Journal* 91 (focus issue): 800-819.
- Ford, C. E. 2004. Contingency and units in interaction. *Discourse Studies* 6//1: 27-52.
- Ford, C. E., B. A. Fox, and S. A. Thompson 1996. Practices in the construction of turns: The "TCU" revisited. *Pragmatics* 6/3: 427-454.
- Garfinkel, H. 1967. *Studies in Ethnomethodology*. Englewood Cliffs, NJ: Prentice Hall.

- Gass, S. M. 1988. Integrating research areas: a framework for second language studies. *Applied Linguistics* 9/2: 198-217.
- Gass, S. M. 1998. Apples and Oranges: or, why apples are not orange and why they don't need to be. *The Modern Language Journal* 82:1 83-90.
- Gass, S. M. 2003. Input and interaction. In: Doughty, C. J. and M. H. Long (eds.): *Handbook in Second Language Acquisition*. Malden, MA: Blackwell.
- Gass, S. M., J. Lee, and R. Roots. 2007. Firth and Wagner 1997: New ideas or a new articulation? *The Modern Language Journal* 91 (focus issue): 788-799.
- Gass, S. and A. Mackey. 2002. Frequency effects and second language acquisition. A complex picture? *Studies in Second Language Acquisition* 24: 249-260.
- Gass, S. and A. Mackey. 2007. Input, interaction and output in second language acquisition. In: VanPatten, B. and J. Williams (eds.): *Theories in Second Language Acquisition. An introduction*. Mahwah, NJ.: Lawrence Erlbaum.
- Gass, S. M. and L. Selinker. 2001. *Second Language Acquisition – An Introductory Course*. Mahwah, NJ: Lawrence Erlbaum.
- Gass, S. M. and E. Varonis 1989. Incorporated repairs in NNS discourse. In: Eisenstein, M. (ed.): *The Dynamic Interlanguage*. New York: Plenum Press.
- Gass, S.M. and E. Varonis. 1994. Input, interaction and second language production. *Studies in Second Language Acquisition* 16: 283-302.
- Gasser, M. and L.B. Smith 1998. Learning adjectives and nouns. A connectionist account. In: *Language and Cognitive Processes* 13/2-3: 269-306.
- Gitsaki, C. 1999. *Second Language Lexical Acquisition. A Study of the Development of Collocational Knowledge*. San Francisco: International Scholars Publications.
- Givón, T. 1985. Function, structure, and language acquisition. In: D. Slobin (ed.): *The Crosslinguistic Study of Language Acquisition*. Hillsdale, NJ: Lawrence Erlbaum.
- Givón, T. 1995. *Functionalism and Grammar*. Amsterdam: John Benjamins.
- Glahn, E., G. Håkansson, B. Hammarberg, A. Holmen, A. Hvenekilde, and K. Lund. 2001. processability in Scandinavian second language acquisition. In: *Studies in Second Language Acquisition* 23: 389-416.
- Goffman, E. 1981. *Forms of Talk*. Blackwell, Oxford.
- Goldberg, A. 1995. *Constructions. A Construction Grammar Approach to Argument Structure*. Chicago: The University of Chicago Press.

- Goldberg, A. 1998. Patterns of experience in patterns of language. In: Tomasello, M. (ed.): *The New Psychology of Language*. Mahwah, New Jersey: Lawrence Erlbaum
- Goldberg, A. 2003. Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences* 7/5: 219-224.
- Goodwin, M. H. and C. Goodwin 1986. Gesture and co-participation in the activity of searching for a word. *Semiotica* 62: 51-75.
- Goulden, R., P. Nation, and J. Read 1990. How large can a receptive vocabulary be? *Applied Linguistics* 11/4: 341-363.
- Granger, S. 1998. Prefabricated patterns in advanced EFL writing: collocations and formulae. In: Cowie, A. P. (ed.): *Phraseology. Theory, analysis, and applications*. Oxford: Clarendon Press.
- Gregg, K. R. 1990. The variable competence model of second language acquisition, and why It isn't. *Applied Linguistics* 11/4: 364-393.
- Gregg, K. R. 2001. Learnability and SLA theory. In: Robinson, P. (ed.): *Cognition and Second Language Instruction* Cambridge: Cambridge University Press.
- Gumperz, J. J. and D. Hymes 1972. Preface. In: Gumperz, J. J. and D. Hymes (eds.): *Directions in Sociolinguistics: The ethnography of Communication*. New York: Holt, Rinehart and Winston.
- Haastrup, K. and B. Henriksen 2000. Vocabulary acquisition: acquiring depth of knowledge through network building. *International Journal of Applied Linguistics* 10/2: 221-239.
- Haastrup, K. and B. Henriksen. 1998. Vocabulary acquisition: From partial to precise understanding. In: Haastrup, K. and Å. Viberg (eds.) *Perspectives on Second Language Vocabulary Research* Lund: Lund Universitet.
- Håkansson, G. 2005. Similarities and differences in L1 and L2 development. In: Pienemann, M. (ed.): *Cross-linguistic aspects of processability-theory*. Amsterdam: John Benjamins.
- Håkansson, G. and U. Nettelbladt 1993. Developmental sequences in L1 (normal and impaired) and L2 acquisition of Swedish. *International Journal of Applied Linguistics* 3: 131-157
- Hakuta, K. 1974. Prefabricated patterns and the emergence of structure in second language acquisition. *Language Learning* 24: 287-297.
- Hall, J. K. 1993. The role of oral practices in the accomplishment of our everyday lives: The sociocultural dimension of interaction with implications for the learning of another language. *Applied Linguistics* 14/ 2: 145-166.
- Hall, J. K. 1995. (Re)creating our world with words: A sociohistorical perspective of face-to-face interaction. *Applied Linguistics* 16/ 2: 206-232.

- Hall, J. K. 1997 A consideration of SLA as a theory of practice: A response to Firth and Wagner. *The Modern Language Journal* 81/3: 301-306.
- Hall, J. K. 2004. Language learning as an interactional Achievement. *The Modern Language Journal* 88/4: 607-612.
- Hall, J. K. 2004. Practicing speaking in Spanish: Lessons from a high school foreign language classroom. In: Boxer, D. (ed.): *Studying Speaking to Inform Second Language Learning*. Clevedon: Multilingual Matters. [referenced in ESK1]
- Hall, J. K. and L. S. Verplaetse 2000. The development of second and Foreign language learning through Classroom interaction. In: Hall, J. K. and L. S. Verplaetse (eds.): *Second and Foreign Language Learning through Classroom Interaction*. Mahwah, NJ: Lawrence Erlbaum
- Harder, P. 2001: The lexico-syntactic symbiosis in a functional perspective. *Nordic Journal of Linguistics* 24: 232-240.
- Hatch, E. 1978. Acquisition of syntax in a second language. In: Richards, J. C. (ed.): *Understanding Second and Foreign Language Learning*. Rowley, Mass. Newbury House.
- He, A. W. 2004. CA for SLA: Arguments from the Chinese language classroom. *The Modern Language Journal* 88/4: 568-582
- Hellermann, J. 2006. Classroom interactive practices for developing L2 literacy: A microethnographic study of two beginning adult learners of English. *Applied Linguistics* 27/3: 377-404.
- Hellermann, J. 2007. The development of practices for action in classroom dyadic interaction: focus on task openings. *The Modern Language Journal* 91/1: 83-96
- Hellermann, J. and E. Cole. forthc. Practices for social interaction in the language learning classroom: disengagements from dyadic task interaction.
- Henriksen, B. 1999a: Three dimensions of vocabulary development. *Studies in Second Language Acquisition* 1999, Special Issue on "Incidental L2 vocabulary acquisition: Theory, current research, and instructional implications" 21/2: 303-317.
- Henriksen, B: 1999b: Ordforråd og ordforrådsindlæring. [vocabulary and vocabulary acquisition] In: Holmen, A. and K. Lund (eds.): *Studier i dansk som andetsprog [studies in Danish as a second language]* Copenhagen: Akademisk forlag.
- Henriksen, B. and K. Haastrup 2000. Describing foreign language learners' productive vocabulary use and development. In: Plag, I. and K.P. Schneider (eds.): *Language Use, Language Acquisition and Language History* Trier: Wissenschaftlicher Verlag Trier.

- Heritage, J. 1984. A change-of-state token and aspects of its sequential placement. In: Atkinson, J. M. and J. Heritage (eds.). *Structures of social action: studies in conversation analysis* Cambridge: Cambridge University Press.
- Hoey, M. 2007. Grammatical creativity: a corpus perspective. In: Hoey, M., M. Mahlberg, M. Stubbs and W. Teubert. (eds.): *Text, Discourse and Corpora*. London: Continuum.
- Hopper, P. J. 1988. Emergent grammar and the a priori grammar postulate. In: Tannen, D. (ed.): *Linguistics in Context*. Norwood: Ablex.
- Hopper, P. J. 1991: Functional explanation in linguistics and the origins of language. *Language and Communication* 11/1-2: 45-47.
- Hopper, P. J. 1998. Emergent grammar. In: Tomasello, M. (ed.): *The New Psychology of Language, volume 1*. Mahwah, NJ.: Lawrence Erlbaum.
- Houtkoop-Stenstra, H. 1994. De interactionele functie van zacht spreken in interviews [the interactional function of low volume in interviews]. *Gramma/TTT* 3: 183-202
- Howarth, P. 1998. Phraseology and second language proficiency. *Applied Linguistics* 19/1: 24-44.
- Hultstijn, J. H. 2001. Intentional and incidental second language vocabulary learning: A reappraisal of Elaboration, rehearsal, and automaticity. In: Robinson, P. (ed.): *Cognition and Second Language Instruction* Cambridge: Cambridge University Press.
- Hunston, S. and G. Francis. 2000. *Pattern Grammar. A Corpus Driven Approach to the Lexical Grammar of English*. Amsterdam: John Benjamins.
- Hutchby, I. and Wooffitt, R. 1998 *Conversation analysis*. Cambridge: Polity
- Hyltenstam, K. 1977. Implicational patterns in interlanguage system variation. *Language Learning* 27/2: 383-411
- Hyltenstam, K. 1985. L2 learners' variable output and language teaching. In: K. Hyltenstam and M. Pienemann (eds.): *Modelling and Assessing Second Language Acquisition* Clevedon: Multilingual Matters.
- Hyltenstam, K. and M. Pienemann 1985. Introduction. In: K. Hyltenstam and M. Pienemann (eds.): *Modelling and Assessing Second Language Acquisition*. Clevedon: Multilingual Matters.
- Hymes, D. 1967. Models of the interaction of language and social setting. *Journal of Social Issues* 23/2: 8-38.
- Hymes, D. 1972. On Communicative competence. In: Pride, J. B. and J. Holmes (eds.): *Sociolinguistics*. Hammondsworth: Penguin
- Hymes, D. 1989. Postscript. *Applied Linguistics* 10/2: 244-252.

- Irujo, S. 1993. Steering clear: Avoidance in the production of idioms. *International Review of Applied Linguistics in Language Teaching* 31/3: 205-219.
- Israel, M. C. Johnson, and P. J. Brooks. 2000. From states to events: The acquisition of English passive participles. *Cognitive Linguistics* 11-1/2: 103-29.
- Jackendoff, R. 2002. *Foundations of Language*. Oxford: Oxford University Press.
- Jordan, G. 2004. Explanatory adequacy and theories of second language acquisition. *Applied Linguistics* 25/4: 538-542.
- Kanagy, R. 1999. Interactional routines as a mechanism for L2 acquisition and socialization in an immersion context. *Journal of Pragmatics* 31: 1467-1492.
- Kasper, G. 1997. 'A' stands for Acquisition. A response to Firth and Wagner. *The Modern Language Journal* 81/3: 307-312.
- Kasper, G. 2004. Participant orientation in German conversation-for-learning. *The Modern Language Journal* 88/4: 551-567.
- Kasper, G. 2007. Conversation-analytic perspectives on language development. Discussant paper presented at AAAL, Costa Mesa, CA. April 21<sup>st</sup>-24<sup>th</sup>, 2007.
- Kasper, G. and E. Kellerman 1997. Introduction: Approaches to Communication Strategies. In: G. Kasper and E. Kellerman (eds.): *Communication Strategies*. Harlow: Longman.
- Kecskes, I. 2003. *Situation-Bound Utterances in L1 and L2*. Berlin: Mouton de Gruyter.
- Kjellmer, G. 1991. A Mint of Phrases. In: Aijmer, K. and B. Altenberg (eds.): *English Corpus Linguistics* (Longman)
- Koshik, I. 2003. Wh-questions used as challenges. *Discourse Studies* 5/1: 51-77
- Kramsch, C. 1986. From language proficiency to interactional competence. *The Modern Language Journal* 70/4: 366-372.
- Krashen, S. 1978. Second language acquisition. In: Dingwall, W. (ed.): *A Survey of Linguistic Science*. Stamford, Conn.: Greylock.
- Krashen, S. D. and R. Scarcella. 1978. On routines and patterns in language acquisition and performance. *Language Learning* 28: 283-300.
- Lafford, B. A. 2007. Second language acquisition reconceptualized? The impact of Firth and Wagner 1997. *The Modern Language Journal* 91 (focus issue): 735-756.
- Lakoff, G. 1991. Cognitive versus generative linguistics: How commitments Influence results. *Language and Communication* 11/1-2: 53-62.

- Langacker, R. W. 1987. *Foundations of Cognitive Grammar, vol. 1. Theoretical prerequisites*. Stanford: Stanford University Press.
- Langacker, R. W. 1988. A Usage-based model. In: Rudzka-Ostyn, B. (ed.): *Topics in Cognitive Linguistics*. Amsterdam: John Benjamins.
- Langacker, R. W. 1991 *Concept, Image, Symbol*. Berlin: Mouton de Gruyter
- Langacker, R. W. 2000. A dynamic usage-based model. In: Barlow, M. and S. Kemmer (eds.). 2000. *Usage-Based Models of Language*. Stanford: CSLI.
- Lantolf, J. P. 1996. SLA theory building: Letting all the flowers bloom! *Language Learning* 46/4: 713-749.
- Lantolf, J.P. 2000. Introducing sociocultural theory. In: J.P. Lantolf (ed.): *Sociocultural Theory and Second Language Learning*. Oxford: Oxford University Press.
- Lantolf, J.P. 2005. Sociocultural second language learning research: An exegesis. In: E. Hinkel (ed.): *Handbook of Research in Second language Teaching and Learning*. Mahwah, NJ: Lawrence Erlbaum.
- Lantolf, J. P. 2007. Sociocultural source of thinking and its relevance for second language acquisition. In: *Bilingualism: Language and Cognition* 10/1: 31–33.
- Lantolf, J. P. and S. L. Thorne. 2006. *Sociocultural Theory and the Genesis of Second Language Development*. Oxford: Oxford University Press.
- Larsen-Freeman, D. and M. Long. 1991. *A Introduction to Second Language Acquisition Research* Harlow: Longman.
- Larsen-Freeman, D. 1991. Second language acquisition research: Staking out the territory. *TESOL Quarterly* 25/2: 315-350.
- Larsen-Freeman, D. 1997. Chaos/complexity Science and second language acquisition. *Applied Linguistics* 18/2: 141-165.
- Larsen-Freeman, D. 2002. Language acquisition and use from a chaos/complexity theory perspective. In: Kramsch, C. (ed.): *Language Acquisition and Language Socialization*. London: Continuum.
- Larsen-Freeman, D. 2004. CA for SLA? It all depends... . *The Modern Language Journal* 88/4: 603-607.
- Larsen-Freeman, D. 2006. The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics* 27/4: 590-619.
- Larsen-Freeman, D. 2007. Reflecting on the Cognitive–social debate in second language acquisition. *The Modern Language Journal* 91 (focus issue): 773-787.



- Larsen-Freeman, D. and M. Long 1991. *An Introduction to SLA Research* Harlow: Longman.
- Laufer, B. 1995. Beyond 2000: A measure of productive lexicon in a second language. In: L. Eubank, L. Selinker, M. Sharwood Smith (eds.): *The Current State of Interlanguage* Amsterdam: John Benjamins.
- Laufer, B. 1998. The development of passive and active vocabulary in a second language: Same or different? *Applied Linguistics* 19/2: 255-271.
- Laufer, B. and P. Nation 1999. A vocabulary size test of controlled productive ability. *Language Testing* 16/1: 33-51.
- Lave, J., and Wenger, E. 1991. *Situated Learning: Legitimate Peripheral Participation*. New York: Cambridge University Press.
- Lerner, G. H. 1991. On the syntax of sentences-in-progress. *Language in Society* 20: 441-458.
- Lewis, M. 1993. *The Lexical Approach* UK: Language Teaching Publications.
- Liddicoat, A. 1997. Interaction, social structure, and second language use: A response to Firth and Wagner. *The Modern Language Journal* 81/3: 313-317
- Lieven, E. V. M., Pine, J. M. and Baldwin, G. 1997. Lexically-based learning and early grammatical development. *Journal of Child Language* 24: 187-219.
- Lieven, E., H. Behrens, J. Speares, and M. Tomasello. 2003. Early syntactic creativity: a usage-based approach. *Journal of Child Language* 30: 333-370.
- Lightbown, P. and N. Spada 1993. *How Languages are Learned*. Oxford: Oxford University Press.
- Liu, E.T.K. and P.M. Shaw 2001. Investigating learner vocabulary: A possible approach to looking at EFL/ESL learners' qualitative knowledge of the word. *International Review of Applied Linguistics in Language Teaching* 39: 171-194.
- Long, M. H. 1983. Native speaker/non-native speaker conversation and negotiation of comprehensible input. *Applied Linguistics* 4/2: 126-141.
- Long, M. H. 1985. Input and second language acquisition theory. In: Gass, S. M (ed.): *Input in Second Language Acquisition*. Rowley, Mass.: Newbury House.
- Long, M. H. 1990. The Least a second language acquisition theory Needs to explain. *TESOL Quarterly* 24/4: 649-666.
- Long, M. H. 1993. Assessment strategies for SLA theories. *Applied Linguistics* 14/3: 225-249.
- Long, M. H. 1997. Construct validity in SLA research: A response to Firth and Wagner. *The Modern Language Journal* 81/3: 318-323.

- Long, M. H. 2007. *Problems in SLA*. New York: Lawrence Erlbaum.
- Loschky, L. 1994. Comprehensible input and second language acquisition. What is the relationship? *Studies in Second Language Acquisition* 16: 303-324.
- Lund, K. 1996. Communicative function and language-specific structure in second language acquisition – a discussion of natural sequences of acquisition. In: E. Engberg-Pedersen, M. Fortescue, P. Harder, L. Heltoft, L.F. Jakobsen (eds.): *Content, Expression, and Structure. Studies in Danish Functional Grammar*. Amsterdam: John Benjamins.
- Lund, K. 1997. Lærer alle sprog på samme måde? [*does everybody learn Danish in the same way?*] Copenhagen: Specialpædagogisk forlag
- Mackey, A. 1999. Input, interaction and second language development. *Studies in Second Language Acquisition* 21: 557-587.
- Mackey, A. 2007. Introduction. The role of conversational interaction in second language acquisition. In: Mackey, A. (ed.): *Conversational Interaction in Second Language Acquisition*. Oxford: Oxford University Press.
- MacWhinney, B. 2000. Connectionism and language learning. In: Barlow, M. and S. Kemmer (eds.): *Usage-Based Models of Language*. Stanford: CSLI
- MacWhinney, B. 2001. The competition model. In: Robinson, P. (ed.): *Cognition and Second Language Instruction* Cambridge: Cambridge University Press.
- MacWhinney, B. 2004. A multiple processes solution to the logical problem of language acquisition. *Journal of Child Language* 31: 883-914.
- MacWhinney, B. 2006. Emergentism – use often and with care. *Applied Linguistics* 27/4: 729-740.
- Markee, N. 2000. *Conversation Analysis*. Mahwah, NJ: Erlbaum.
- Markee, N. and G. Kasper 2004. Classroom talks: An introduction *The Modern Language Journal* 88/4: 491-500.
- McLaughlin, B. 1990. Restructuring. *Applied Linguistics* 11/2: 113-128.
- McLaughlin, B., T. Rossman, and B. McLeod 1983. Second language learning: An information-processing perspective. *Language Learning* 33/2: 135-158.
- McNamara, T. F. 1995. Modelling performance: Opening Pandora's Box. *Applied Linguistics* 16/2: 159-179.
- Meara, P. 1996. The dimensions of lexical competence. In: Brown, Gillian, Kirsten Malmkjær and John Williams (eds.): *Performance and Competence in Second Language Acquisition* Cambridge: Cambridge University Press.

- Meisel, J., H. Clahsen, and M. Pienemann. 1981. On determining developmental stages in natural second language acquisition. *Studies in Second Language Acquisition* 3/2: 109-135.
- Melka Teichroew, F. J. 1982. Receptive versus productive vocabulary: A survey. *Interlanguage Studies Bulletin* 6/2: 5-33.
- Melka, F. 1997. Receptive vs. productive aspects of vocabulary. In: Schmitt, N. and M. McCarthy *Vocabulary: Description, acquisition, and pedagogy*. Cambridge: Cambridge University Press.
- Mellow, J. D. 2006. The emergence of second language syntax: A case study of the acquisition of relative clauses. *Applied Linguistics* 27/4: 645-670.
- Miller, G. A. 2003. The cognitive revolution: a historical perspective. *Trends in Cognitive Sciences* 7/3: 141-144.
- Mitchell, R. and F. Myles. 1998/2004. *Second Language Learning Theories*. London: Arnold (1<sup>st</sup>/2<sup>nd</sup> ed.)
- Mondada, L. and S. Pekarek Doehler. 2004. Second language acquisition as situated practice: Task accomplishment in the French second language classroom. *The Modern Language Journal* 88/4: 501-518.
- Moon, R. 1997. Vocabulary connections: multiword items in English. In: Schmitt, N. and M. McCarthy (eds.): *Vocabulary: Description, acquisition, and pedagogy*. Cambridge: Cambridge University Press.
- Mori, J. 2004. Negotiating sequential boundaries and learning opportunities: A case from a Japanese language classroom. *The Modern Language Journal* 88/4: 536-550.
- Mori, J. 2007. Border crossings? Exploring the intersections of second language acquisition, conversation analysis, and foreign language pedagogy. *The Modern Language Journal* 91 (focus issue): 849-862.
- Myles, F., J. Hooper, and R Mitchell. 1998. Rote or rule? Exploring the role of formulaic language in classroom foreign language learning. *Language Learning* 48/3: 323-363.
- Myles, F., R. Mitchell, and J. Hooper. 1999. Interrogative chunks in French L2. A basis for creative construction? *Studies in Second Language Acquisition* 21: 49-80.
- Nation, I.S.P. 1990. *Teaching and Learning Vocabulary* Boston, Mass: Heinle and Heinle.
- Nattinger, J. 1980. A lexical phrase grammar for ESL. *TESOL Quarterly* 14/3: 337-344.
- Nattinger, J. 1988. Some current trends in vocabulary teaching. In: Carter, R. and M. McCarthy (eds.): *Vocabulary and Language Teaching* Harlow: Longman.
- Nattinger, J. P. and J. S. DeCarrico. 1992. *Lexical Phrases and Language Teaching*. Oxford:

Oxford University Press

- Nesselhauf, N. 2003. The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics* 24/2. 223-242
- Nesselhauf, N. 2005. *Collocations in a Learner Corpus*. Amsterdam: John Benjamins.
- Newmeyer, F. J. 1991. Functional explanation in linguistics and the origins of language. *Language and Communication* 11/1-2: 3-28
- Newmeyer, F. J. 1998. *Language Form and Language Function*. Cambridge, MA: MIT Press.
- Ohta, A. S. 2001. *Second Language Acquisition Processes in the Classroom*. Mahwah, NJ.: Lawrence Erlbaum.
- Paribakht, T. 1985. Strategic competence and language proficiency. *Applied Linguistics* 6/2: 132-146.
- Pawley, A. and F. H. Syder 1983. Two puzzles for linguistic theory. In: Richards, J. C. and R. W. Schmidt (eds.): *Language and Communication*. Harlow: Longman.
- Peters, A. M. 1983. *Units of Language Acquisition*. Cambridge: Cambridge University Press.
- Pienemann, M. 1985. Learnability and syllabus construction. In: Hyltenstam, K and M. Pienemann (eds.): *Modelling and Assessing Second Language Acquisition*. Clevedon: Multilingual Matters.
- Pienemann, M. 2002. Issues in second language acquisition and language processing. *Second Language Research* 18/3:189–192.
- Pienemann, M. 2003. Language processing capacity. In: *The Handbook of Second Language Acquisition*. Ed. C. J. Doughty and M. H. Long. Oxford: Blackwell.
- Pienemann, M. 2007. Processability theory. In: VanPatten, B. and J. Williams (eds.): *Theories in Second Language Acquisition. An introduction*. Mahwah, NJ. Lawrence Erlbaum.
- Pienemann, M. and G. Håkansson. 1999. A unified approach toward the development of Swedish as L2. *Studies in Second Language Acquisition* 21: 383-420.
- Pienemann, M., M. Johnston, and G. Brindley. 1988. Constructing an acquisition based procedure for second language assessment. *Studies in Second Language Acquisition* 10/2: 217-243.
- Pinker, S. 1994. *The Language Instinct* London: Penguin.
- Pinker, S. 1999. *Words and Rules* London: Weidenfeld and Nicolson.
- Polio, C. and S. Gass. 1998. The role of interaction in native speaker comprehension of non-native speaker speech. *The Modern Language Journal* 82: 308-319.

- Poullisse, N. 1993. A theoretical account of lexical communication Strategies. In: Schreuder, R. and B. Weltens (eds.): *The Bilingual Lexicon* Amsterdam: John Benjamins.
- Poullisse, N. 1997. Some words in defense of the psycholinguistic approach: A response to Firth and Wagner. *The Modern Language Journal* 81/3: 324-328
- Poulsen S. 2005. *Collocations as a language resource. A cognitive and functional study in English phraseology*. Unpublished PhD-dissertation, University of Southern Denmark. URL: [http://www.sdu.dk/Information\\_til/Studerende\\_ved\\_SDU/Din\\_uddannelse/Phd\\_humaniora/phd\\_afhandling\\_iframe.aspx](http://www.sdu.dk/Information_til/Studerende_ved_SDU/Din_uddannelse/Phd_humaniora/phd_afhandling_iframe.aspx)
- Radden, G 1992. The Cognitive Approach to natural language. In: Pütz, M. (ed.): *Thirty Years of Linguistic Evolution*. Amsterdam: John Benjamins.
- Read, J. 1988. Measuring the vocabulary knowledge of second language learners. *RELC Journal* 19: 12-25
- Read, J. 1997. Assessing vocabulary in a second language. In: Clapham, C. and D. Corson: *Encyclopaedia of Language and Education, Volume 7: Language Testing and Assessment*. Dordrecht: Kluwer Academic Publishers.
- Read, J. 2004. Research in teaching vocabulary. *Annual Review of Applied Linguistics* 24: 146-161.
- Read, J. and P. Nation. 2004. Measurement of formulaic sequences. In: Schmitt, N.. (ed.) 2004. *Formulaic Sequences*. Amsterdam: John Benjamins.
- Reder, S. 2005. The "Lab School". *Focus on Basics* 8a. 1-7. URL: [http://www.ncsall.net/fileadmin/resources/fob/2005/fob\\_8a.pdf](http://www.ncsall.net/fileadmin/resources/fob/2005/fob_8a.pdf)
- Reder, S., K. A. Harris, and K. Setzler. 2003. A Multimedia Adult Learner Corpus. *TESOL Quarterly* 37/3: 546-557.
- Renouf, A. and J. McH. Sinclair 1991. Collocational frameworks in English. In: Aijmer, K. and B. Altenberg (eds.): *English Corpus Linguistics*. Harlow: Longman.
- Revier, R. L. 2005. *Phraseological Competence: Empirical study of collocations in written L2 and L1 production by Danish learners of English*. Unpublished master's thesis, University of Copenhagen.
- Riley, P. 1996. Developmental sociolinguistics and the competence/performance distinction. In: Brown, G., K. Malmkjær and J. Williams (eds.): *Performance and Competence in Second Language Acquisition* Cambridge: Cambridge University Press.
- Ritchie, W. C. and T.K. Bhatia (eds.). 1996. *Handbook of Second Language Acquisition*. San Diego, CA: Academic Press.

- Rumelhart, D. and J. McClelland. 1987. Learning the past tenses of English verbs: implicit rules or parallel distributed processes? In: MacWhinney, B. (ed.): *Mechanisms of Language Acquisition*. Hillsdale, NJ. Lawrence Erlbaum.
- Sacks, H. E. A. Schegloff, and G. Jefferson 1974. A Simplest systematic for the organization of turn-taking for conversation. *Language* 50: 696-735.
- Sato, C. 1986. Conversation and interlanguage development: rethinking the connection. In: Day, R. (ed.): *Talking to Learn: Conversation in Second Language Acquisition*. Rowley, Mass.: Newbury House.
- Schachter, J. 1990. Communicative competence revisited. In: Harley, B., P. Allen, J. Cummins, and M. Swain (eds.): *The Development of Second Language Proficiency* Cambridge: Cambridge University Press.
- Schegloff, E. A. 1968. Sequencing in conversational openings. *American Anthropologist* 70: 1075-1095.
- Schegloff, E. A. 1979. The relevance of repair to syntax-for-conversation. In: Givón, T. (ed.): *Syntax and Semantics: Discourse and Syntax* (vol. 12). New York: Academic Press.
- Schegloff, E. A. 1988. Discourse as an interactional Achievement II: An exercise in conversation analysis. In: D. Tannen (ed.): *Linguistics in Context: Connecting Observation and Understanding*. Norwood, NJ: Ablex.
- Schegloff, E. A. 1992. Repair after next turn: The last structurally provided defense of intersubjectivity. *The American Journal of Sociology* 97/ 5: 1295-1345.
- Schegloff, E. A. 1996. Turn organization: one intersection of grammar and interaction. In: Ochs, E., E. A. Schegloff, and S. A. Thompson (eds.): *Interaction and Grammar*. Cambridge: Cambridge University Press.
- Schegloff, E. A. 2007. *Sequence Organization in Interaction*. Cambridge: Cambridge University Press.
- Schegloff, E.A., I. Koshik, S. Jacoby, and D. Olsher. 2002. Conversation analysis and applied linguistics. *Annual Review of Applied Linguistics* 22: 3-31.
- Schegloff, E. A., E. Ochs, and S. A. Thompson. 1996. Introduction. in: Ochs, E., E. A. Schegloff, and S. A. Thompson (eds.): *Interaction and Grammar* Cambridge, Cambridge University Press 1-51.
- Schieffelin, B, and E. Ochs. 1986. Language socialization. *Annual Review of Anthropology* 15: 163-191.
- Schmidt, R. 1983. Interaction, acculturation, and the acquisition of communicative competence: A case study of an adult. In: Wolfson, N. and E. Judd (eds.): *Sociolinguistics and Language Acquisition*. Rowley, Mass.: Newbury House.

- Schmidt, R. 1990. The role of consciousness in second language learning. *Applied Linguistics* 11/2: 129-158.
- Schmidt, R. 2001. Attention. In: Robinson, P. (ed.): *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press
- Schmitt, N. 2000. *Vocabulary in Language Teaching*. Cambridge: Cambridge University Press.
- Schmitt, N. (ed.) 2004. *Formulaic sequences*. Amsterdam: John Benjamins.
- Schmitt, N. and R. Carter. 2000. Lexical phrases in language learning. *The Language Teacher*.  
URL: <http://www.jalt-publications.org/tlt/articles/2000/08/schmitt>
- Schmitt, N. and R. Carter. 2004. Formulaic sequences in action: an introduction. In: Schmitt, N.. (ed.) 2004. *Formulaic sequences*. Amsterdam: John Benjamins.
- Schmitt, N., Z. Dörnyei, S. Adolphs, and V. Durow. 2004. Knowledge and acquisition of formulaic sequences. A longitudinal study. In: Schmitt, N.. (ed.) 2004. *Formulaic sequences*. Amsterdam: John Benjamins.
- Schmitt, N., S. Grandage, and S. Adolphs. 2004. Are corpus-derived recurrent clusters psycholinguistically valid? In: Schmitt, N.. (ed.) 2004. *Formulaic sequences*. Amsterdam: John Benjamins.
- Schmitt, N., and G. Underwood. 2004. Exploring the processing of formulaic sequences through a self-paced reading task. In: Schmitt, N.. (ed.) 2004. *Formulaic sequences*. Amsterdam: John Benjamins.
- Searle, J. 1969. *Speech Acts*. Cambridge: Cambridge University Press.
- Selinker, L. 1972. Interlanguage. *International Review of Applied Linguistics in Language Teaching* 10: 209-231
- Selinker, L. 1992. *Rediscovering Interlanguage* Harlow: Longman.
- Selinker, L. and D. Douglas 1985. Wrestling with 'context' in interlanguage theory. *Applied Linguistics* 6/2: 190-204.
- Selting, M. 2005. Syntax and prosody as methods for the construction and identification of turn-constructional units in conversation. In: Hakulinen, A. and M. Selting (eds): *Syntax and Lexis in Conversation: Studies on the Use of Linguistic Resources in Talk-in-Interaction*. Philadelphia: John Benjamins.
- Selting, M.. 2000. The construction of units in conversational talk. *Language in Society* 29: 477-517.

- Sfard, A. 1998. On two metaphors for learning and the dangers of choosing just one. *Educational Researcher* 27/2: 4-13.
- Sharwood Smith, M. 1986. Comprehension versus acquisition: Two ways of processing input. *Applied Linguistics* 7/3: 239-256.
- Sharwood-Smith, M. 1991. IL, conceptual confusions and new beginnings. In: Phillipson, R., E. Kellerman, L. Selinker, M. Sharwood Smith and M. Swain (eds.): *Foreign/Second Language Pedagogy Research*. Clevedon, Avon: Multilingual Matters.
- Shohamy, E. 1996. Competence and performance in language testing. In: Brown, G., K. Malmkjær and J. Williams (eds.): *Performance and Competence in Second Language Acquisition* Cambridge: Cambridge University Press.
- Siewierska, A. 1992. Layers in FG and GB. In: Fortescue, M, P. Harder, L. Kristoffersen (eds.): *Layered Structure and Reference in a Functional Perspective*. Amsterdam: John Benjamins.
- Sinclair, J. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sinclair, J. and A. Renouf. 1988. A lexical syllabus for language learning. In: Carter, R. and M. McCarthy (eds.): *Vocabulary and Language Teaching*. Harlow: Longman.
- Skehan, P. 1998. *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.
- Stefanowitsch, A. and S. Th. Gries. 2003. Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics* 8/2: 209-243.
- Stubbs, M. 1995. Collocations and cultural connotations of common words. *Linguistics and Education* 7: 379-390.
- Stubbs, M. 2007. Quantitative data on multi-word sequences in English: the case of the word *word*. In: Hoey, M., M. Mahlberg, M. Stubbs and W. Teubert. (eds.): *Text, Discourse and Corpora*. London: Continuum.
- Swain, M. 1985. Communicative competence: Some rules of comprehensible input and comprehensible output in its development. In: S.M. Gass and C.G. Madden (eds.): *Input in Second Language Acquisition*. Cambridge: Newbury House.
- Tarone, E. 1983. On the variability of interlanguage Systems. *Applied Linguistics* 4/2: 142-164.
- Tarone, E. 1990. On variation in interlanguage: A response to Gregg. *Applied Linguistics* 11/4: 392-400.
- Tarone, E. 2000. Still Wrestling with 'context' in interlanguage theory. *Annual Review of Applied Linguistics* 20:182-198.
- Tarone, E. 2002. Frequency effects, noticing, and creativity. Factors in a variationist interlanguage framework. *Studies in Second Language Acquisition* 24: 287-296.



- Taylor, D. S. 1988. The meaning and use of the term "competence" in linguistics and applied linguistics. *Applied Linguistics* 9/2: 148-168.
- Taylor, J. R. 1995. *Linguistic Categorization*. Oxford: Oxford University Press.
- Taylor, J. R. 1998. Syntactic constructions as prototype categories. In: Tomasello, M. (ed.): *The New Psychology of Language*. Mahwah, New Jersey: Lawrence Erlbaum.
- Taylor, J. R. 2002. *Cognitive Grammar*. Oxford: Oxford University Press.
- Ten Have, P. 1999. *Doing conversation analysis*. London: Sage.
- Thelen, E. and E. Bates. 2003. Connectionism and dynamic systems: are they really different? *Developmental Science* 6/4: 378–391.
- Tomasello, M. 1998: Introduction: A Cognitive-Functional Perspective on language Structure. In: Tomasello, M. (ed.): *The New Psychology of Language*. Mahwah, New Jersey: Lawrence Erlbaum.
- Tomasello, M. 2000. First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics* 11/1-2: 61-82.
- Tomasello, M. 2003. *Constructing a Language*. Cambridge: Cambridge University Press.
- Tomasello, M. 2003. Introduction: Some surprises for Psychologists. In: Tomasello, M. (ed.): *The New Psychology of Language - volume 2*. Mahwah, New Jersey: Lawrence Erlbaum.
- Tummers, J, K. Heylen, and D. Geeraerts. 2005. Usage-based approaches in cognitive linguistics: A technical state of the art. *Corpus Linguistics and Linguistic Theory* 1-2: 225-261.
- Underwood, G., N. Schmitt, and A. Galpin. 2004. The eyes have it: An eye movement study into the processing of formulaic sequences. In: Schmitt, N.. (ed.) 2004. *Formulaic sequences*. Amsterdam: John Benjamins.
- Unsworth, S., T. Parodi, A. Sorace, and M.-Y. Scholten. 2006. Introduction. In: Unsworth, S. (ed): *Paths of development in L1 and L2 acquisition: in honour of Bonnie D Schwartz*. Philadelphia, PA: John Benjamins.
- Van Lancker-Sidtis, D. and G. Rallon. 2004. Tracking the incidence of formulaic expressions in everyday speech: Methods for classification and verification. *Language and Communication* 24: 207-240.
- Van Lier, L. 1996. *Interaction in the Language Curriculum. Awareness, Autonomy and Authenticity*. Harlow: Longman.
- Van Lier, L. 2002. An ecological-semiotic perspective on language and linguistics. In: Kramsch, C. (ed.): *Language Acquisition and Language Socialization*. London: Continuum.

- Van Valin, R.D. 1991: Functionalist linguistic theory and language acquisition. *First Language* 11: 7-40.
- VanPatten, B. 1990. Attending to form and content in the input. An experiment in consciousness. *Studies in Second Language Acquisition* 12: 287-301.
- VanPatten, B. 2007. Input processing in adult SLA. In: VanPatten, B. and J. Williams (eds.): *Theories in Second Language Acquisition. An introduction*. Mahwah, NJ.: Lawrence Erlbaum.
- VanPatten, B. and J. Williams (eds.) 2007. *Theories in Second Language Acquisition. An introduction*. Mahwah, NJ.: Lawrence Erlbaum
- Waara, R. 2004. Construal, convention, and constructions in L2 speech. In: Achard, M. and S. Niemeier (eds.): *Cognitive linguistics, second language acquisition and foreign language pedagogy*. Berlin: Mouton de Gruyter.
- Wagner, J. 2004: The Classroom and Beyond. *The Modern Language Journal* 88/4: 612-616.
- Wagner, J. and Gardner, R. 2004. Introduction. In: Gardner, R. and J. Wagner (eds.): *Second Language Conversations*. London: Continuum.
- Watson-Gegeo, K. 2004. Mind language and epistemology: towards a language socialization paradigm for SLA. *The Modern language journal* 88/3: 331-350.
- Watson-Gegeo, K. and S. Nielsen. 2003. Language socialization in SLA. In: Doughty, C. and M. Long (eds.): *Handbook of Second Language Acquisition*. Malden, MA: Blackwell
- Weinert, R. 1995. The role of formulaic language in second language acquisition: A review. *Applied Linguistics* 16/2:180-205.
- Wenger, E. 1998. *Communities of Practice. Learning, Meaning, and Identity*. Cambridge: Cambridge University Press.
- Wesche, M. and T. S. Paribakht. 1996. Assessing second language vocabulary knowledge: depth versus breadth. *The Canadian Modern Language Review* 53/1: 13-40.
- White, L. 2003. On the nature of interlanguage representation: Universal grammar in the second language. In: Doughty, C. and M. Long (eds.): *Handbook of Second Language Acquisition*. Malden, MA: Blackwell.
- White, L. 2007. Linguistic theory, universal grammar, and second language acquisition. In: VanPatten, B. and J. Williams (eds.): *Theories in Second Language Acquisition. An introduction*. Mahwah, NJ.: Lawrence Erlbaum.
- Widdowson, H.G. 1989.: Knowledge of language and Ability for use. *Applied Linguistics* 10/2: 128-137

- Wood, D. 2002. Formulaic language in acquisition and production: implications for teaching. *TESL Canada Journal* 20/1: 1-15.
- Wray, A. 1999. Formulaic language in learners and native Speakers. *Language Teaching* 32: 213-231.
- Wray, A. 2000. Formulaic sequences in second language teaching: Principle and practice. *Applied Linguistics* 21/4: 463-489.
- Wray, A. 2002. *Formulaic Language and the Lexicon*. New York, NY: Cambridge University Press.
- Wray, A. and M. R. Perkins. 2000. The functions of formulaic language: an integrated model. *Language and Communication* 20: 1-28.
- Young, R. F. and E. R. Miller. 2004. Learning as changing participation: discourse roles in ESL writing conferences. *The Modern Language Journal* 88/4: 519-535.
- Zobl, H. and J. Liceras 1994. Functional Categories and acquisition Orders. In: *Language Learning* 44: 159-180.
- Zuengler, J and K. Cole. 2005. Language socialization and second language learning. In: Hinkel, E. (ed.): *Handbook of Research in Second Language Teaching and Learning*. Mahwah, NJ: Lawrence Erlbaum.
- Zuengler, J. and E. R. Miller 2006. Cognitive and sociocultural perspectives: two parallel SLA worlds? *TESOL Quarterly* 40-1: 35-58.



**Appendices:** Inventory counts and traced utterances, initial and final recording periods (ESK2)

Appendix 1: inventory count and traced utterances, Valerio, recording period 1 (ESK2).

**Bold types indicate verbatim multi-word repetitions.**

Target utterance:	closest matches:	utterance schema + <i>operation(s)</i>
1 Big here is big	is good, no good	is /it's x <i>Substitution + add-on.</i>
<b>2 My wife she's</b> the market one melon	<b>My wife she</b> shopping for me now	my wife she <i>substitutions (several, number uncertain; topic-comment)</i>
3 <b>No is</b> melon	<b>no is</b> for profession	no is x <i>substitution + drop (meaning extension)</i>
4 Orange () here	here is big	is / it's x <i>substitution + rearrange + drop</i>
5 <b>It's good</b>	she check for <b>is good no good</b>	is / it's x (meaning extension)
6 It's ... <b>No is</b> different for red 6a it's no different	<b>no is:</b> see above input rep	no is x <i>Substitution + insert / add-on</i>
7 In color	in euro, dollars	in + abstract <i>substitution</i>
8-9 How do you say + input rep (2)	<b>what do you say (3)</b>	<b>MWE</b> <i>add-on</i>
10 <b>It's</b> very different	<b>it's:</b> see above	it's / is x <i>substitution + insert</i>
11 in the sentence ((+ reading)) no is at noon is at afternoon <b>in</b> my home	<b>in</b> my home	in + location ( <i>two</i> ) <i>substitution(s)</i> ; no is x, it's / is x <i>two substitutions</i>

12 In my country is_	<b>in my home</b>	in + location, <i>substitution</i>
13 In my country it's afternoon	<b>in my country</b> (input rep <sup>96</sup> ) <b>it's good</b>	in my x + it's / is x <i>two substitutions</i>
14 The united state people he like the noon	<b>I like gold</b>	I, he like x <i>Two substitutions + add-on (topic-comment)</i>
15 <b>In my country</b>	<b>In my country</b>	<b>MWE</b> (in + location)
16 Two or three	evening or night	x or y <i>Two substitutions</i>
17 two to three	n/a	
18 <b>This here</b>	<b>this here</b>	<b>MWE</b>
-- But in the us the people like lunch () the noon	reading own writing	but (add-on) // in + loc (substitution) // x like y (two substitutions) // the noon (add-on; verbatim rep) <i>Combination of (at least) two schemas; three substitutions, two add-ons</i>
19 She makes		input rep.
20 This () is okay?	<b>This here</b> (MWE) <b>It's okay</b>	it's / is x <i>add-on (first use of it's okay as interrogative; combination of previously used material to achieve new meaning.)</i>
21 <b>It's okay?</b>	<b>It's okay</b>	<b>MWE</b>
22 I like eat	<b>I like gold</b>	x like y, <i>substitution</i>

---

<sup>96</sup> We have *in my home* which is thought to overrule the input rep category. However because it is afforded by the material at hand, we do not defined it as a MWE.

23-24 No is good (2)	<b>no is</b> different	no is x, <i>substitution</i> ; second instantiation is in same turn, not considered a MWE
25-26 In this country (2)	<b>in my country</b>	in + location, <i>substitution</i> ; second instantiation is in same turn, not considered a MWE
27 I no understand	no is x I don't know (MWE)	no- negation. <i>Combination of previously used material. First instantiation of I no verb.</i>
28 <b>What do you say</b> (followed by input reps)	<b>what do you say (3)</b>	<b>MWE</b>
29 The food is take one and two hours		copula + verb, <i>lexis afforded</i>
30 Five minutes it takes me five minutes		input rep.
31 Is for the Nordstrom () here	What <b>is for the</b> people you <b>For the</b> years	it's / is x + for NP, <i>two substitutions + add-on</i> . Combination of schemas.
32 Is for you, her	is for: see above	it's x + for NP <i>two substitutions - combination of schemas</i>
33 Just working	<b>just</b> for the years Quit my <b>work</b> (2) My wife she shopping <b>Moving</b> for me please?	<i>combination of previously used material</i>
34 I am no forget	<b>I am</b> go I <b>no</b> understand	no-negation, copula + verb. combination of two schemas, substitution.
35 You is teach level d	cop + verb (extension of <i>I am verb</i> )	cop + verb. <i>three substitutions.</i>

36 <b>What happen</b> for the test	<b>what happen</b>	<b>MWE +</b> for NP <i>substitution</i>
37 I am forget for the last name and first name you	<b>I am</b> no <b>forget</b> for NP	copula + verb, for NP, <i>Combination of two utterance schemas with an additive and inserted, adding on yet another NP to the for NP schema. Number of required operations uncertain.</i>
38 I don't forget this	<b>I don't</b> know	do-negation, <i>substitution + add-on</i>
39 Oh my goodness	n/a	
40 The restaurant for Indian food	I am go <b>the restaurant</b> <b>for:</b> see above	<i>combination of previously used elements or previously used element added on to for NP</i>
41 Yes I do, I do	practise.	Input rep.
42 I like the movies the dancing	<b>I like</b> gold	x like y <i>substitution. Number of operations uncertain.</i>
43 Do you dancing tango?	I like the movies the <b>dancing</b> What <b>do you</b> say (MWE) <b>Do you</b> repeat please	do you verb? <i>substitution; number of operations uncertain</i>
44 For clothes	<b>for:</b> see above	for NP, <i>substitution</i>
45 Always clothes	n/a	
46 no always is	is / it's x (see above)	is / it's x <i>add-on</i>
47 <b>What do you say</b> no always clothes?	<b>What do you say</b> for anniversary married, have xxx, 'estuvo'	<b>MWE +</b> <i>substitutions, inserts; number uncertain</i>



48 No is always	<b>no is</b> x / substitute	<b>no is</b> x = substitute / no always is = <i>rearrange</i>
49 Do you like (∅), Lorena?	What <b>do you</b> say	do you verb? <i>drop + substitution + add-on.</i>
50 Is for you	<b>that is for you</b>	<i>verbatim rep + drop + possible meaning extension could also be two utterance schemas it's x and for NP</i>
51 For <b>do you like</b> to movie or xxx	<b>do you like</b>	<i>add-on; + MWE + uncertain number of add-ons.</i>
52 Why <b>do you like</b> (2)	<b>do you like</b>	<i>add-on + MWE</i>
53 Why you like	why do you like	<i>add-on + MWE + drop</i>
54 Bless you	n/a	
55 Your mom she makes bread every day	input rep	notice the topic-comment structure
56 Do you like movies the horror	<b>do you like</b> <b>I like the movies the dancing</b>	<b>MWE;</b> Uncertain number of substitutions
57 I don't know		MWE
58 I don't know what happened	<b>I don't know</b> <b>What happen (2)</b>	combination of two MWEs
59 I have a small party?	n/a	
60 <b>I don't know</b>		<b>MWE</b>
61 Winnie Pooh is for my baby (2)	that <b>is for</b> you	<i>x is for y; two substitutions, the latter one may hold a substitution in itself.</i>

62 Reading no writing	<b>x-ing</b> three dollars five () no fifteen	x no y; x-ing; two substitutions
63-66 What is (4)	practise!	Input rep
67 From my house	where are you <b>from</b> <b>my</b> x (frequently employed)	combination of previously used material uncertain number and kinds of operations
68 For the house	<b>for:</b> see above	for NP, substitution
-- Xxx house is for xxx and xxx	inaudible	
69 for the MAX xxx is for five minutes	copula pattern <b>for:</b> see above	for NP + for NP, <i>two substitutions</i> , <i>combined by copula</i>
70 Who goes to Bally's	I go for the park	Interaction affordance
71 <b>I don't know</b> I ask the question for he	<b>I don't know</b>	combination of previously used material: <b>MWE</b> , potential utterance schema <i>ask the question</i> , (drop <i>for</i> , <i>add-on</i> I) + utterance schema <i>for x</i> . There is no schema for combining all these elements.
72 Five seconds	number + NP	five x; <i>substitution</i> (or <i>number x</i> ; <i>two substitutions</i> )
73 It takes five minutes	input rep (see note 1 in text)	number x + <i>add-on</i> (afforded)
74 For me tired	<b>for me</b> I like gold	<b>MWE</b> + add-on
75 Five minutes tired	five minutes	number x + <i>add-on</i>
total: 76 audible multi-word utterances (tokens; 75-76 inconsistency: note 6a+b above)		

**token count:**

**MWUs produced by way of**

recycled utterance schemas	55	72,4 %
input reps	11	14,4 %
combinations of previously used...	5	6,6 %
untraceable (n/a)	5	6,6 %
	<u>76</u>	<u>100,0 %</u>

**preliminary utterance schema summary of inventory:**

- 1 It's / is x
- 2 No is x
- 3 In + abstracts
- 4 For + NP<sup>97</sup>
- 5 In + location
- 6 I like NP → x like y
- 7 No-negation
- 8 Do-negation
- 9 Copula + verb
- 10 Do you verb<sup>98</sup>
- 11 My wife she x (topic-comment)
- 12 x or y (summer 2003) → is x or is y (autumn 2003) *hypothesis: this pattern is often used for meta-linguistic purposes, e.g. to ask about 'correctness': 'can I or can you'?*
- 13 x is for y (closely related to *for NP*)
- 14 x no y (closely related to *x or y*)
- 15 x-ing
- 16 my x
- 17 number x

---

<sup>97</sup> This pattern usually means 'in terms of x' or 'as far as x is concerned'.

<sup>98</sup> Do you like the movies the horror, I like the movies the dancing...!

Appendix 2: inventory count and traced utterances, Valerio, recording period 6 (ESK2).

**Bold types indicate verbatim multi-word repetitions.**

<u>Target utterance</u>	<u>traced matches:</u>	<u>utterance schema, <i>operation(s)</i></u>
2 I can say he doesn't or doesn't he (2)	<b>I can</b> use, pay	I can V, <i>substitution</i> x or y, <i>substitutions</i> combination of two schemas
3 And <b>this one that's okay</b>	you can take <b>this one</b> <b>This one</b> is difficult <b>That's okay</b>	Combination of two <b>MWEs</b> ; <i>and</i> add-on
4 Is a complint sentence	<b>is a</b> public university?	Is a x; <i>two substitutions</i>
5 Will be is won't he	copula	Copula combination + input reps
6 And the three what did you have	<b>what do you have</b> <b>did</b> practice july, sept 03, apr, jul 04 <b>what did you</b> write (2), say (2), do (3) what did he say	blend of two schemas; what did you verb, and do you have; add-ons
7 Oh my gosh		MWE
8 you can check exactly what is first and what is second	<b>you can</b> change, use you <b>check</b> x <b>What is</b> x (frequent) ...talk about <b>what is</b> ... or <b>what is...</b> you don't know <b>what is</b> legal	<i>combination of 3 utterance schema tokens, 2 utterance schema types with insertion of and you can verb; what is x</i>
9 Oh my gosh		MWE

<b>10 I don't know</b>		<b>MWE</b>
11 Thank you		<b>MWE</b>
12 You can say ((writes))	see above	you can v, <i>substitution</i>
<b>13 It's easy</b>		<b>MWE</b>
14 It's the short name <b>for this one</b> ( <i>short addition</i> )	<b>it's</b> x it's a small one, it's a squirrel, it's a small family, similar <b>for this one</b>	<i>combination of 2 utterance schemas. The it's x schema, however, may need to be subdivided into several schemas; e.g., here it might be the it's the x-schema. The closest match, however, is probably more than two substitutions away.</i>
15 It's the_ little <b>difficult</b>	<b>it's little</b> better, confused it's new / great experience for you	it's x. combination of previously....
16 yeah is long	<b>is</b> x	utterance schema, add yeah
<b>17 You're welcome</b>		<b>MWE</b>
18 That's here	<b>that's</b> x <b>Here:</b> frequently employed	that's x; substitution
-- Yeah okay because I used xxx	object inaudible.	
<b>19 Thank you</b>		<b>MWE</b>
<b>20 What is</b> keens (=kinds)	<b>what is</b> x	utterance schema, add-on (input rep)
<b>21 You ask me</b>	the first <b>you ask me</b>	<b>verbatim rep</b>

22 [Driving to school is more difficult than take the bus] for the school

**for the x.** *To get at for, a highly frequent item, it's necessary to not only trace previous uses, but also to get at different meanings; in this case it's to.*

For NP; Utterance schema, substitution

**23-24 Yeah exactly (2)**

**MWEs**

Welcome

25 Maybe in a house is more rel no not relax  
uh is more uhm *I don't mean relax* **what is** (gest)  
Comfortable to live in apartment (co-constr)

**in + loc; is more x; is more x; I don't v; what is x;** input rep

**Really**

26 Yeah opposite

yeah x

**27 It's stressful**

**it's x.** *this shows the usefulness and the power of the patterns under investigation. Here we have a brand new item (as far as we know; he seems to pick it up from the teacher in this session) which easily employs in an existing pattern.*

**28 What is your question**

**what is your** work

verbatim rep; also utterance schema what is your x

**29 I don't know the name** two beers

**MWE,** add-ons

**30 I don't know**

**MWE**

**31 I don't know names** beers

**MWE,** add-on

32 Is more stressful

**is more** slowly

utterance schema, substitution

33 What is xxx (=of two options what is more x?)

**what is** more difficult

utterance schema, substitution

**34 Is better?**

*has not necessarily been used as a question before **verbatim rep***

35 - 36 Where do you live (2)

*we have where are you from for example, where do you x, and do you x, and on numerous occasions we have live. So again we have a new combination of items already in use. In may we also see where did you live as practiced form and PS, and in June 05 there's an interaction between Virg and partner which features the issue of living and in which where you live is afforded.*

**Where do you** shop, go, buy? Utterance schema, substitution

37 You know how driving from Gresham here?

**Do you know...**

**Drive – driving** used before

**From – here:** for the Nordstrom here. *We've established that for sometimes means from – and so in this case learning has been empirically substantiated to involve traversing a path from NNL to NL. It does however seem to be a relatively infrequent learning process.*

*You know how* has never been used before – the question is then if this pattern is traceable or not. It's afforded but not previously used.

Do you know + add-on.

38 Beaverton is I think one hour driving for the 26<sup>th</sup> ave

**I think** normal  
**for NP**

copula, **MWE**, one hour x-ing (subst),  
for NP (subst)

**I think** transportation

**I think** (Ø, NL)

**I think** yes because you don't pay

**One hour x-ing:** practise 27 jul 04

**Driving:** frequent item

**39 Because** it's <spn>

**because it's** very cheap

Maybe it's own schema; otherwise it's *it's x*  
+ add-on

**40 It's more** is stressful driving to school

see above

it's more, is x, input rep

41 or you can take the bus	<b>you can take</b> this one <b>take the bus</b>	combination of two MWEs, + add-on
42 What time you come <b>in the morning</b> ?	<b>What time:</b> in previous usage (NNL) When did <b>you come</b> to the US (3 oct 03 practised form) <b>come</b> to class: practiced form earlier this session! <b>You come</b> back you no tell me? <b>In + time</b> and <b>in + place!!!</b>	Combination of three verbatim reps. However, <i>you come</i> could also be a variant of do you verb? + drop.
43 Go outside your home? (=leave your home)	you use aeroplane for <b>go</b> china? <b>Outside your home:</b> 21 jul 05. in this instance, we see a previously established pattern used in a NNL way. If we deduce, by implication of the fact that time has passed, in the previous instances that learning has taken place in that Valerio seems to have traversed a path going from NNL to NL, then, surely, we must deduce that the opposite has happened here; ‘learning’ has implied going from NL to NNL! Another navigable road towards solving this anomaly involves taking a different turn altogether in approaching the basic issue of what it means to learn a language. But which one???	
44 <b>It's the same</b> me (here?)	in <b>the same</b> house They talk <b>the same</b> Is the same (2)	MWE + add-on
45 – 46 You see the traffic? (2) (=do you...)	<b>(do) you</b> x?	utterance schema, substitution (drop)
47 I think you don't have traffic	<b>I think:</b> see above <b>you don't have</b> accidents? <b>I don't have, you don't</b> read, understand, ø	<b>MWE, verbatim rep,</b> substitution
48 <b>You can take</b> 26 <sup>th</sup> or		<b>MWE/verbatim rep,</b> x or y





<b>58 Is only</b> one for all people everybody <b>in the kuhneh</b> (=class, while laughing)	<b>Is/it's x</b> <b>only one</b> <b>For x, for all</b> the world, the questions, day	blend of an utterance schema and a MWE + utterance schema + insert/add-on + utterance schema
<b>59 Where do you live</b>		<b>verbatim rep</b>
60 No <b>it's stressful it's stressful</b>		<b>local MWE</b> , add-on
61 <b>It's stressful</b> because it I live in Beaverton	<b>because:</b> see above	local <b>MWE, because</b> , I live in Mexico/Beaverton/Portland
62 <b>You can take 26 is I think one hour driving</b>	<b>I live in</b> Mexico	<b>combination of verbatim reps / MWEs</b>
63 because the traffic is slowl	here <b>is</b> more <b>slowly</b>	is more x
64 no I like more take (the?) bus	<b>I like</b> gold: initial use. <b>More:</b> seems to be most frequent in <i>one more (x)</i> . <b>I like</b> eat is the closest match (which might have been part of a task). <b>Take</b> x is another pattern in use, especially in terms of transport.	Combination of previously used....
65 <b>what is you</b> (=your note)	<b>what is your</b> question?	What is your x, substitution
<u>66 depend</u> on the traffic	<i>depend on = untraceable!</i>	
67 or depend on where <b>you work</b>	when <b>you work</b> do <b>you work</b> here () and where <b>where</b> they check	(x) or y, depend, + combination of
previously...		
68 <b>is more</b> better	<b>is more</b> x	subst

69 yeah exactly

MWE

70 you can go wherever place and more easy yeah

**can** see above & previous  
**go** see above & previous  
**more** see above & previous  
wherever: lexical addition.

*You can go // go + loc // and // more x //  
add-on*

71 it's the people receive money for months  
but they're not working

**money for** the government  
they borrow **money**  
**they're** doing  
when **they is** come in here  
but **they** don't talk  
**they** talk the same  
**they** live together  
**they** have place  
receive: new item (substitution, borrow money)

*it's x // the people receive money // for x //  
but // neg*

72 when **you is** olders

73 **when you is** after 75 7

**when you V** (see above)

74 you can **receive** one pension

**you can:** see above

you can verb, substitution(s)

75 I know

MWE

76 ĩn the future

(rep input) – but the schema is there: in + time, place, abstract (all in one here1)

77 **That's** yeah for the rest of your life (!!!)

how is **your life** different... *untraceable*

78 Yeah exactly

MWE

79 Your child		
80 Your children (lexical)		combination of previously....
81 But depend		
82 Depends what your are writing (= in the will)	<b>depend:</b> see above Talk about <b>what</b> is... <b>You're writing</b>	combination of previously...
83 Or my moneys 55 my son 55 my son	x or y, <b>my</b> x	two schemas, substitution
84 Or my daughter <b>take</b> a 100%	x or y, my x, <b>take</b> x	three schemas, substitution
85 Hello Helen <u>how are you</u>	<b>how you</b> doing?	uncertain category = untraceable.
86 Good morning	n/a	uncertain category = untraceable.
87 He can come y sit with you	<i>He can come + sit + with you</i>	x can verb; verb with you (item: live, stay) three substitutions
88 I don't remember the name	<b>I don't remember</b>	<b>MWE + verbatim rep</b> , add-on
89 What is the name		MWE
90 I don't remember		MWE
<b>91 You can sit with</b> Helen		you can verb, verb with x; three substitutions
92 Helen she is single	<b>single:</b> <b>she is</b> American	most often used in nativelylike way, in the sense that it opposes 'living with somebody'. Here meaning is broadened to conform to 'being alone', to a peer missing a partner for the task. Parallel with 'you born single' (not a twin) (27 jul 04). Topic-comment structure.

**93 It's different**

**94 Because is** exactly house for

**95 depends** you have money  
for these guys they have money

96 it's perfect // for // because / it's cheap // and // they have / **all things** // you need / for **the first time** // for one baby time // it's perfect

**MWE**

(because) it's x, uncertain number of substitutions;  
**you have x, for** [people], **they have x**, three substitutions. The only thing is *guys*: V uses it very infrequently; but it is quite often afforded.

it's x, because it's x, two substitutions (*it's cheap* is a verbatim rep, however probably not frequent enough to be a MWE. They have, all things, you need, for, the first time, are all verbatim reps. *one baby time*, however, seems to be creative (and also not conform to a purist view on grammaticality.) FOR NP: very elaborate schema by now!

**97 it's okay**

98 only this is one thing I don't like about xxx (=oriented to as America)

**MWE**

only // this is / one thing // I don't like about xxx; add-on; MWE; combination of two frequent items; neg; add-on

only: very frequent  
**this is**  
the other thing  
the thing for the plats and cups  
**I don't like** + add-on

99 it's little

verbatim rep

100 it's little thing	<b>it's x</b>	verbatim rep + add-on
101 and you don't read it?	<b>You don't v</b>	add-on – neg.
102 Yeah sure		yeah x

total: 102 audible Multi Word Utterances

**token count:**

MWUs by way of		
recycled utterance schemas	92	90,2 %
input reps	1	1,0 %
combinations of previously used...	5	4,9 %
untraceable	4	3,9 %
	<u>102</u>	<u>100,0 %</u>

preliminary utterance schema summary of inventory:

- 1 I, you, he can verb
- 2 x or y
- 3 is a / it's a x / it's the
- 4 copula
- 5 what did you Verb
- 6 what is x
- 7 For NP → issue here probably semantic (for this one = possible MWE)
- 8 That's x
- 9 You ask me (verbatim rep, should be counted to establish if it's a MWE)
- 10 In + loc
- 11 Is more x / more x
- 12 I, you don't v
- 13 Yeah x
- 14 What is your question (verbatim rep; utterance schema what is your x)
- 15 Where do you x

- 16 (do) you know (how)
- 17 from – here
- 18 Because it's x
- 19 What time verbatim rep, semantised
- 20 (do) you verb
- 21 when you verb
- 22 what (did) he say?
- 23 For playing the sport (for + input rep)
- 24 I ask you verbatim rep
- 25 Only x
- 26 I, he change for
- 27 I live + loc
- 28 I like (ø/x)
- 29 Depend on the traffic / where you work
- 30 *go + loc / dir*
- 31 *it's x // the people receive money // for x // but // neg prog*
- 32 you is
- 33 for the rest of your life
- 34 your x
- 35 what you verb
- 36 my x
- 37 how are you
- 38 verb with you
- 39 She is adj
- 40 prn have np
- 41 you need verbatim rep
- 42 This is x

Appendix 3: inventory count and traced utterances, Carlos, recording period 1 (ESK2).

**Bold types indicate verbatim multi-word repetitions.**

Target utterance:	closest matches:	utterance schema + <i>operation(s)</i>
1. Eh this one ()	<b>This one</b>	<b>MWE</b>
2. like this okay?	Like that <b>This okay</b>	combination of schemas; like x (substitution) and <b>MWE</b>
3. I like	new construction, NNL	input rep (practised form).
4. Have you engineer (NNL; engineer = afforded by material)	you do not have address? You no have Have a good day!	you no verb; rearrange + drop + substitute
5. You no like ice skating?	You no have?	you no verb? Substitution ( <i>like x</i> is practised item)
6. Like that?	<b>Like that?</b>	<b>MWE</b>
7. Hot tea		co-construction / input rep
8. Yeah maybe	both items frequently employed	combination of previously used material ( <i>yeah x</i> might be a schema)
9. Maybe like it cold	<i>maybe</i> : frequent item	



	Like x (see above)	<i>Like x</i> (practiced item) substitution, add-on + add-on ( <i>maybe x</i> might be a schema; is reused, see below)
10. No ready because the film () develop → <i>ready, develop</i> : new items	because she work in a factory Because the question is... It's no correct	it's no x, drop + substitute; because x v, two substitutions
11. It's in thirty-six picture (not <i>it's on...?</i> )	It's in the Broadway	it's in location; substitution
12. I have only twenty-six	we have only one class I have ehm I have job	I have (only) x (blend of schemas?: I have x / only x?)
13. Maybe ready for the twenty-nine for the last class	Maybe like it cold he need a help for the write We have only one class for the test You need work for the money last name	maybe x (see above; substitution); for NP (substation); for NP (substitution)
14. Maybe the twenty-nine I have already (all ready / it ready??)	I have: see above	maybe x (substitution); I have x (substitution + add-on)
15. Don't worry don't worry	no worry	<b>MWE</b> <i>modified to a native-like variety</i>
16. I give you the picture	give you coming [directions]	give you x (substitution + add-on)
17. can you help me? (writes on the board, struggling to spell 'thirty')	can you write, spell, see...	can you x (substitution, ad-on)

18. what is much better? (i.e., what does it mean in Spanish?)	what is your first name, this,	what is x (substitution) (first time used in this sense (use of pattern shared by Valerio!))
19. I worked	she works at 8 o clock	x work (substitution)
20. Yeah thanksgiving work	yeah maybe (see above)	yeah x (substitution); x work (drop);
21. Wait wait wait wait when?	Wait wait wait it's no correct	wait n times + x (substitute)
22. I stay at home	she work in <b>at home</b>	untraceable verb pattern; stay + <b>MWE</b>
23. I no cook nothing	I no remember	I no verb (substitute)
xxx sleepy	sleep, sleeping	<i>inaudible, not counted</i>
24. Because my cousin () wife they go in the California	because: see above In the: see above I go home (go + dir) He go the run She go open You go write the uj	because x v (uncertain number of substitutions) go + dir / in + location (substitution)
25. Sometimes maybe the thanksgiving together in Vancouver	only sometimes (3) sometimes I smoke	sometimes x, maybe x, in + location (uncertain number of substitutions)

26. But no they this time yeah (partial co-constr)	but, no, yeah , they: elements used previously this one	Combination of previously used elements (this x)
27. Only () in the night	not write only read in the: used frequently. this is first use in relation to time.	only x; in + time
28. In the night	verbatim repetition	
29. For the dance	for the test	For NP (substitution)
30. I was uh dancing (where did you go?)	untraceable	
31. I think uh first avenue	I think so	I think so (drop, add-on)
<b>32. Thank you</b>		
33. Thank you	<b>thank you</b>	<b>MWEs</b>
34. Do you is the people has one pee or two pee	where do you live Do you need a help Friends or students He has different You no have Number x	do you X (repair); x have y; x or y (uncertain number of substitutions), inexplicable (but not untraceable) copula.
35. You know	<b>D'you know?</b>	<b>MWE</b> (drop)
36. Yes but eh my question is	one question	combination of previously known material (/ perhaps verbatim rep of previous input; prior to this, the teacher says <i>my question is</i> )

37. How many pee have the word people	how many you have	how many x / x have y (rearrange, substitute; number of operations uncertain)
38. Can you spell people for me ('spell' given by teacher)	can you: see above	Can you verb; spell x; for NP ( <b>for me</b> : possible <b>MWE</b> )
39. Only one?		
40. One two?	Only sometimes, one class (see above)	only x ( <b>only one</b> : possible <b>MWE</b> )
41. Oh for the next year	for the money (see above)	for NP (substitution)
42. come in here	come on man over here	come + direction (two substitutions)
43. He like watch the baseball (co-constr)	I like (see above)	x like y + input rep (counted as utterance schema recycling because of the co-constructed nature of the utterance)

Total: 43 audible multi-word utterances

token count:

MWUs produced by way of:

untraceable	2	4,65 %
input reps	2	4,65 %
combinations of previously used...	2	4,65 %
recycled utterance schemas	37	86,05 %
	<u>43</u>	<u>100,00 %</u>

preliminary utterance schema summary of inventory:

1 it's no x

- 2 because x v,
- 3 it's in location
- 4 x have y
- 5 only x?
- 6 maybe x
- 7 for NP
- 8 give you x
- 9 can you x
- 10 what is x
- 11 x work
- 12 yeah x
- 13 wait n times + x
- 14 I no verb
- 15 go + direction
- 16 in + location
- 17 sometimes x
- 18 this x
- 19 do you X
- 20 x or y
- 21 how many x
- 22 spell x
- 23 come + direction
- 24 number x

Appendix 4: inventory count and traced utterances, Carlos, recording period 4 (ESK2).

**Bold types indicate verbatim multi-word repetitions.**

<u>Target utterances</u>	<u>closest match</u>	<u>schemas + operations.</u>
1 what did you do yesterday	how <b>did you</b> pronounce this word?	do / did you verb + various add-ons
2 oh this is about the weather	<b>This is</b> the evening it's <b>about</b> today	this is x; about x, substitutions
3 that's nice	<b>that's</b> good	that's x, substitution
4 but eh yesterday I went to the to beaverton xxx	I, he <b>went to the</b> shopping, (practise) practise <b>went</b> (interesting <i>I went to shopping</i> )	I went to the x, substitution + add-on
5 terrible this happen now → <i>terrible</i> : new item	what <b>happen</b> (with you) this, now: previously used	combination of previously used elements + add-on
<b>6 you know</b> the in iraq eh the people no like but eh was <b>a lot of</b> bombs on the streets in the I talk to <b>the people</b> outside Something you <b>don't like</b> ? I <b>don't like</b> the raining What <b>was</b> her name <b>On the street</b>		you know // in loc // the people // don't like // but // was <b>a lot of NP // on the streets</b> (combination of utterance schemas and MWEs; multiple substitutions)
7 yeah that's true	<b>that's</b> adj: see above	yeah x; that's x, substitution
8 the man with the mask xxx	I don't need a <b>mask</b> I go <b>with</b> this one	combination of previously used elements (perhaps

rearrangement?) + with x (also, **this one** has been previously established to be a **MWE**)

9 oh you didn't see

**see** this one?  
Teacher how **did** you pronounce this?

do/did you x + negation, substitution

10 yeah I saw that in the on the TV

Check prev  
21 Oct 03: practise **saw**  
feb 04: **I saw** snow  
watch **the tv**  
**in, on** x *it seems to be the first combination of prep and TV*

x see y; in / on loc, multiple substitutions

11 man with a mask he has the four people **in front of** him with the

**I have the** microphone  
We **have the** big party

with x; x have y; number x; in front of x; with x (multiple substitutions)

**12 I don't know how you say that**

**I don't know how you say that**

**MWE**

13 They was talking about how was the accident or\_ was **thinking about**\_ there **a lot of people** was thinking she didn't die

What do you **think about** this?  
Don't worry **about** it  
Let me **think about** it  
**Think:** generally frequent  
**Talking about:** april 02; heavy input jan 04  
**How:** I don't know how  
Was: see above

→ *accident*: possible new item

x copula talking about // how // copula NP // thinking about // a lot of people // cop thinking // x do neg y (combination of utterance schemas and **MWEs**; multiple substitutions)

<b>14 I don't know</b>	<b>I don't know</b>	<b>MWE</b>
15 Wow interesting	wow with a boyfriend	wow x, substitution
16 In April of this year	the song <b>of</b> Michael Jackson <b>This</b> one (MWE) <b>This</b> time (MWE) <b>This</b> word, class Last words <b>of</b> the first sentence .	In + time // of // this x (multiple substitutions) <i>Seemingly, this is the only instance of of used in connection with 'time'.</i>
17 Wow her <b>boyfriend</b> is latin guy or	guy, latin: previously used	wow x, her x, copula, multiple substitutions
18 White guy	white, guy, previously used	combination of previously used elements
19 From here	where you <b>from</b> I'm <b>from</b> Mexico From where <b>Here:</b> frequently used	combination of previously used elements
<b>20 How many daughters you have?</b>	how many do you have	<b>how many x, do you have, substitution, MWE</b>
<b>21 How many daughters you have or how many children?</b>		<b>how many x, do you have, substitution, MWE</b>
22 Just one	have <b>just one</b> question	<b>MWE</b>
<b>23 That's good just one</b>	<b>that's good</b> <b>just one</b>	<b>MWEs</b>



- 24 No I don't have any **I don't have any** idea, money MWE / utterance schema, substitution  
 -- The parties xxx **every weekend** **every:** seems to be used a lot with 'time'  
*inaudible, not counted*
- 25 **Every weekend** go to the nightclub to dance  
**Go to** school, supermarket, Mexico  
*every weekend I love to go out to dance to the nightclubs*  
**MWE**, go to loc + inf (perhaps 'to the nightclub' is a MWE)
- 26 Sometimes here in downtown sometimes in Beaverton  
**Here in** Portland  
 Sometimes: previously used Sometimes x; (here) in + loc (multiple substitutions)
- 27 **I don't know** bec\_ salsa **I don't know** **MWE** + add-on
- 28 **I love to dance** **I love to dance** **MWE**
- 29 **In Portland they have** um it's in spain restaurant uh Fernando's hideaway  
**They have um...** MWE? In + loc; x have y; it's in + loc + name (potential MWE; multiple substitutions)
- 30 It's on the first avenue **it's on first avenue** **MWE**
- 31 Yeah because I go to the\_ to dance at f's hideaway and many people from the many states they  
 dance very well like Japanese people and Chinese people  
 Yeah because // I go to the // to dance // at f's hideaway // and // **many people** // from many states  
 // they dance very well // like Japanese and Chinese people //  
 because x; go to inf; many x; from + loc; they dance; very  
 adj; like x; (multiple substitutions)
- 32 They dance very well salsa **I don't know how** uh they understand  
 they dance very well salsa // I don't know how // they understand

**Many** times, things, days, hours, rabbits, more...

**Very** funny, small, good, different, close, easy, fast

June 02: **very well** seems to be understood, and in 2002 it's also an object of study (in relation to 'speak English' questions)

verbatim reps; MWEs; combination they understand

33 The name place is Fernando's

I know this **place**

**The name** the restaurant, of the teacher...

copula: previously used

the name x, substitution

34 It's um the first avenue between yamhill and taylor

See above.

**MWE**

35 After ten

**after** that

after x, substitution

Xxx from eight to ten

**from** wednesday/thursday to saturday

For six to nine!!! 30 sep 04

I work in the evenings five **to ten**

from x to y (time, substitution)

**36 I don't know** who is the teacher **right now** but uh sometimes it's a girl sometimes it's a man

**It's x**

**I know who is** Pedro is

**The teacher**

Sometimes: see above

**MWE I don't know who is** // the teacher // **right now** //  
sometimes x // it's x // it's x (combination of utterance  
schemas, multiple substitutions)

**37 I think** it's uh ten dollars for class

**I think**

**it's** free

five **dollars**

for NP

MWE + it's x; x dollars; for NP (multiple substitutions)

38 no you can still there xxx you don't need to pay again no

**you can v**  
**you need to pay**  
**I don't need to pay**

(at least) two substitutions, three add-ons

39 um go to dance or go to the\_ take the classes?

**Go to take** a break

**Take class (MWE)**, this class, two classes

go + infinitive (2) (substitutions)

40 xxx the classes is um xxx **I think** Thursday for three days Thursday Friday and Saturday

they have classes **for** six to nine

**I think**

copula; for NP (extension of for + time?; *been here for four years*); **MWE**

41 uh (2) **I think** it's\_ the best day go to there like when it's a\_ **a lot people** to dance there xxx on Thursday

**it's** time to go...

**go (to) + loc**

like: previously used on many occasions

**when it's** snowing (2), raining, more frequent: when I/you.

**the best x**; thing (practise), friend

**to dance**

**on** tuesday

**I think, a lot of people, to dance = MWEs**

the best x, when it's x, go to + location, on + day =

utterance schemas, multiple substitutions

42 because um on Fri\_ in Friday and Saturday the people go to the many\_ many places you know

**many x**

**you know** the Mexican, Sylvania

**go to**: see above

**because x**

on + day (see above)

the people (see above)

because x, on + day, the people go + loc., many x **you know (MWE)** (multiple substitutions)

43 yeah **I like** there **I go** there so **I was working there for one year**

I have been here **for four years**

I live in this neighbourhood **for one year** with half

**I was working there**

**I like**

**Go + location**

→ again we have (creative) combinations of previously known elements, overwhelmingly in the form of utterances schemas. It's very hard to quantify this, however, because there's overlaps and fluidity among the identified schemas in a given multi-schema utterance, there's combinations of schemas and MWEs (and overlaps among those, too), there's repairs and pauses etc. Even though this quantification is close to impossible, it's still very obvious that L2 use and learning is still very much based on what we might call developing stability (routinisation).

44 uh I was xxx the\_ (+) prep cook

this was really great, it was really funny, I was working...

x was y (substitutions)

45 went to shopping

**went to shopping**

MWE (see above)

46 he played play the ((gesture))  
the piano and the guitar

**play** music, instruments, in the band...

x play y (substitutions)

-- **I xxx**

inaudible

47 I want to join a band (c-c)

somebody **wants to** sit...?

You **want to** be ø? (nml)

I played in the band

want to v (substitution; 'join a band' afforded)

48 Maybe I need look in the newspaper maybe x

**I need** inf (3)

	In + loc	maybe x, I need infinitive; in + loc; utterance schemas, multiple substitutions
49 On some page	In some group, some band	some x (add-on, <i>on</i> ) <u>or alternatively</u> : on + loc (substitution)
<b>50 I don't know</b> exactly what where the newspaper <b>I need to</b> look		<b>MWE</b> , (I don't know what = entrenched; repair = where) NP; I need infinitive.
51 Maybe I can find <b>some band</b>		maybe x; I can V; some x – utterance schemas, multiple substitutions
51 Many bands in ehm where I'm from	<b>where I live</b> (MWE) <b>I'm from</b> Mexico	many x; in + loc (blend of where I live (MWE) and I'm from Mexico (MWE?))
52 in the city xxx ehm () when it's eh () xxx holidays in ehm where I'm from	in + loc <b>is</b> x when it's x (interesting, <b>when</b> is used more often as a conjunction than an interrogative marker) <b>where I'm from</b>	in + loc, when it's x, <b>where I'm from</b> , two utterance schemas, one possible <b>MWE</b>
53 I like the bands Mexican	<b>I like</b> np <b>Mexican</b> used previously as post-modifier	I like x; x Mexican (two substitutions)
54 Bass guitar and drums		afforded in interaction
55 It's a good thing	<b>it's a</b> secret, nightclub...	utterance schema, substitution

56 I like playing	<b>I like playing (MWE?)</b>	<b>MWE / verbatim recycling</b>
<b>57 Yes I do but</b>	<b>yes I do (a frequently practised utterance in the class)</b>	<b>MWE</b>
58 I want to be a famous to play the music	afforded, but also recycled: I want to repeat, play To play in a band, guitar, music Famous To be	I want to x, substitution to play x, substitution recycled item to be x, substitution
59 I can play any kind of music	<b>any kind of</b> animals I can verb (see above) Play music	any kind of x, substitution I can v, substitution play x (perhaps a MWE, play music)
60 I play the bass guitar and drums	I play x	substitution
61 Blessing is like uh when you (gest) bless you	your question <b>is like</b> this <b>When you</b> 're ready (also, see above on <i>when</i> )	x is like this (substitution, drop); when you v (substitution)
62 When you going home he say god bless	at <b>home (go home)</b> they say okay	when you v; go + dir; x say y (multiple substitution)
63 Is like a holy week (it's)	<b>it's like</b> this	it's like x, substitution (x is like y, two substitutions?)
<b>64 I don't know</b>		<b>MWE</b>

65 I success <b>I don't</b> work or sleep → success: affordance	<b>I don't</b> v	substitution
<b>66 I don't know I don't know</b> man		<b>MWE</b>
67 In my apartment I have a the keyboard	<b>in</b> time vs. place ( <b>in my</b> house) <b>I have</b> np	in my x, substitution I have x, substitution
68 When I am alone <b>over there</b> I can learn	<b>when I</b> call, saw... <b>I can</b> v <b>I'm</b> happy, fine, oldest, not <b>Over there</b>	when I verb substitution I can verb substitution I'm x substitution <b>MWE</b>
69 I put a little music and playing	<b>I put</b> like this, here this <b>a little</b> bit, more x (both MWEs) playing: recycled item	I put x; a little x (two substitutions), insert <i>and</i> + add verb
70 I play the instrumental music	play: see above	play x substitution
71 Play classic music	play: see above	play x substitution
72 By myself                    by myself		<b>MWE</b>
73 I just thinking xxx	I just live I'm just working	I('m) (just) V(-ing) (substitution)
74 But I don't write I don't write the (gest)	<b>I don't</b> v	substitution
75 Yeah I forgot it	<b>forgot:</b> first use of past tense, forget and forgotten used previously	

76 But sometimes I record the music	<b>but sometimes</b> (used previously; both are frequent in other patterns) <b>record:</b> untraceable <b>I play the instrumental music</b> <b>I love the music</b> (recurring)	<i>I verb the music</i> , often used with <i>play</i> as verb, is argued to sanction this utterance (substitute)
77 Because this skateboard have a record xxx	because x this letter, hand, places (x) have x skateboard (means keyboard; repaired by other participant) record: emerges from context (same as above)	because x (substitution); this x (substitution); have x (substitution)
78 You push the uh (motions pushing a button)		untraceable
79 Yeah but <b>I need to</b> have <b>a lot of</b> time to for <b>to do</b> that <b>I need to have</b>		<i>The last part untraceable; counts as such.</i>
80 Put the note xxx right here	<b>put</b> : see above the note: afforded by context <b>Right here</b> (MWE)	put x, substitute
81 This note is not correct I need to	<b>is not correct</b> <b>I need to</b>	this x, substitute + <b>to</b> MWEs / recycled utterances
82 <b>I need to</b> correct this I need to change for <b>another one</b>	correct previously used as verb <b>I need to change</b> <b>I change</b> the Thursday <b>for</b> Friday <b>Another one</b>	I need to x, substitution change x for y, substitutions <b>MWE</b>



83 And you checking the notes there	let me <b>check</b> <b>check</b> how you write there used previously the notes: afforded (context same as above)	x check y (multiple substitutions)
84 It's hard it's hard	<b>it's</b> adj	substitution
<b>85 I need to have a lot of time</b>	<b>I need to have a lot of time</b>	<b>MWE / verbatim rep</b>
86 And they did the notes	did: used previously as both auxiliary and main verb the notes: afforded in the on-going interaction	
87 the book is ehm el notte book	the book: frequently used item copula: frequently use	combination of previously used items
<b>88 I think</b> notte book (2; could be counted as 2)	<b>I think</b> the dishwasher	I think NP, substitution
89 Because in mexico we have some books	because x <b>in + loc</b> <b>x have y</b> some x	substitution substitution substitutions substitution
<b>90 Like this</b>	<b>like this</b>	<b>MWE</b>
91 And the page have the song	some <b>page</b> <b>Have</b> <b>The song</b> of Michael Jackson	combination of previously used elements
92 In mexico they have the word	<b>in + loc</b>	<b>in Mexico (MWE?)</b>

	<b>x have y</b> this <b>word</b>	substitutions
93 In mexico we have <b>like this</b> (gestures, sings)	<b>In mexico we / they...</b> We have x Like this	<b>MWE</b> substitution MWE
94 And then <b>right here</b> we have the notes	and then: frequent in previous interactions, as well as immediately to Carlos's own production in this interaction, as affordance MWE We have x	substitution
95 Okay we can go to the xxx	<b>we can V</b> go to location	substitution substitution
96 You <b>wanna</b> go to the store	<b>go to location</b> <b>want to verb</b>	substitution substitution
<b>97 You don't</b> wanna go to the store / shop?	You don't verb Want to verb Go to location	substitution substitution substitution
98 To a store/shop (nodding)	to + loc	substitution
99 Across from the		verbatim rep
100 Yeah wanna go?	<b>Wanna go</b>	<b>MWE?</b>
101 Why not?		verbatim rep
102 Come on man	<b>come</b> here. (same recipient!)	

	<b>Come on man</b> (what happen with you) verbatim rep / MWE?	
103 where do you put that (1804)	<b>I put here this</b> <b>where do you live</b>	substitution substitution
104 what you <b>think about</b> it (1822)	<b>what do you think...</b>	<b>MWE?</b>
105 I don't have anything (2208)	<b>I don't have any</b>	substitution
106 for what?	<i>For</i> and <i>what</i> previously used separately.	Not recorded as chunked before. Count as <i>untraceable</i> because the two items together here serve a specific purpose in the on-going interaction: partner: you turn; Carlos: for what?
107 can you say that	how do you <b>say that</b> ? Can you verb?	substitute
108 how can I say that	say that: see 107 <b>how can I say?</b>	MWE + add-on; or can I verb + substitute, add-ons
<b>109 Like this</b>	<b>like this</b>	<b>MWE</b>
110 What number?	Number: afforded (also previously used)	what x (kind, question; substitution)
111 Easier for you?	<b>Good for you</b> It's not <b>easy</b> man <b>Easier for you</b> (previous practise)	good for you; substitute – or <b>for you (MWE)</b> + add-on(?)
<b>112 I don't know</b>	<b>I don't know</b>	<b>MWE</b>
113 What does it mean the: (points)	<b>what does it mean</b> this latter what does this say?	MWE / substitution

### What do you think MWE

114 Excuse me	excuse me	MWE
115 I have here eh <b>what does it mean the:</b> darkness ([what's it] mean) no exact match of I have here; here: extremely frequent		<b>MWE</b>  I have x, add-on
116 It's a bad word	bad word: afforded by teacher ('I don't want to use curse words, curse words are very bad') <i>bad</i> and <i>word</i> respectively used by Carlos previously.	it's a x, substitution. The afforded combination of previously known elements are easily incorporated into the <i>it's a x</i> -schema.
117 Bad word	see 116	recycled
118 yeah sometimes I use the bad words when when at the restaurant ((pointing)) yeah x sometimes x use: afforded here by co-participant (also used previously) bad words: see 115-116 when: see various examples above.		Multiple schemas and substitutions, as well as affordances
about you know xxx work ((shakes head))	talk/think about you know discourse marker (see above) <i>inaudible, leave out</i>	
119 about work	afforded by co-participant's question <i>about what?</i>	
-- eihn xxx they say ahr ((gesturing wildly)) xxx		<i>inaudible</i>

120

yeah because he didn't know how to fly () brwch ((plane crash sound))

yeah x

because x

do neg

how to verb

substitution

substitution

substitution

substitution

total: 120 audible multi-word utterances

**token count: 120 MWUs**

**MWUs produced by way of:**

untraceable	3	2,5 %
input reps	0	0 %
combinations of previously used...	9	7,5 %
recycled utterance schemas	108	90 %
	<u>120</u>	<u>100,0 %</u>

**Preliminary utterance schema summary of inventory:**

- 1 do / did you verb
- 2 this is x;
- 3 about x
- 4 it's / that's x
- 5 I went to the x
- 6 yeah x
- 7 do negation
- 8 x see y;
- 9 (here) in / on loc
- 10 with x;
- 11 x have y;
- 12 number x;
- 13 in front of x
- 14 I('m) (just) V(-ing) → x (cop) talking/thinking about
- 15 In + time

- 16 this x
- 17 her x,
- 18 copula / cop neg
- 19 how many x,
- 20 go ((to) (loc)) + inf (embedded to + loc)
- 21 Sometimes x;
- 22 because x;
- 23 many x;
- 24 from + loc;
- 25 very adj;
- 26 (it (is)) like x
- 27 after x
- 28 the name x
- 29 from x to y (time)
- 30 for np
- 31 you can v
- 32 extension of for + time
- 33 the best x
- 34 when it's, you, I...
- 35 I like v
- 36 play music, instruments...
- 37 want to v
- 38 I need inf;
- 39 some x
- 40 maybe x;
- 41 x Mexican
- 42 go + dir;
- 43 x say y
- 44 I put x;
- 45 a little x,
- 46 *I verb the music*
- 47 change x for y
- 48 check x
- 49 where/what/when do/es x (most frequent 'do you' → interactional)

50 what x

51 how to verb

## ENGLISH SUMMARY.

The research presented in this collection of five research articles and three background chapters was originally framed around a wish to explore the role of 'formulaic language' in L2 learning from the perspective of functional-cognitive linguistics. On the basis of empirical evidence, however, the research interest quickly became centred on a wish for exploring the fruitfulness of applying the framework of Usage-Based Linguistics (UBL) to general investigations into L2 learning. The reason for this change was that the question of whether L2 learners start out from formulas and gradually start analysing them to use the individual constituents in other linguistic structures, or if they start from the learning of combinatorial rules with practice ensuring the entrenchment of certain formulas over time, seemed fundamentally misguided and flawed. Rather, formulas and more general patterns were found to co-exist at all points in development, at least as far as the data and focal students investigated here are concerned.

My data come from the Multimedia Adult English Learner Corpus, which consists of audio-visual recordings of classroom interaction in an English as a Second Language classroom in Portland, Oregon. The ESL classrooms, in which the recordings were made, were equipped with six video cameras mounted in the ceiling. Four of those were fixed, and two were moveable by remote control. The two latter cameras each followed a student wearing a wireless microphone; students were given these microphones to wear on a rotational basis. The teacher wore a microphone at all times in the class. The final database of the inquiries in the five research papers consists of transcripts from approx. 70 classroom sessions each consisting of three hours of recordings in which my two focal students, Carlos and Valerio, are either wearing a microphone or sitting next to someone wearing a microphone. Both Carlos and Valerio are Mexican-Spanish speaking learners of English.

What has been invoked as a guiding principle in the research carried out here is the UBL path of learning from formulas via partially schematised patterns to fully abstract constructions. This modus operandi proved extremely useful and empirically valid from the beginning. In the first research paper to apply the UBL framework to longitudinal SLA research, *Are recurring multi-word expressions really syntactic freezes? Second Language Acquisition from the perspective of Usage-Based Linguistics*. (co-authored with T. Cadierno; henceforth EC), we found positive evidence for the UBL learning trajectory, with the system emerging in acquisition characterized as



the gradual abstraction of regularities that link expressions as constructions. *Do-negation* learning was found to be initially heavily reliant on one specific instantiation of the pattern, *I don't know*, with productivity gradually increasing as the underlying knowledge seemed to become increasingly abstract, as reflected in type and token frequencies. The findings suggested that L2 learning is indeed item-based, that expression entrenchment is dependent on token frequency (as in the case of *I don't know*) and that more abstract pattern and construction learning is dependent on type frequency, as expected.

In turn, and chronologically, the five research papers each brought about insights which spawned new research questions. The second paper, *Constructing another Language – Usage-Based Linguistics in Second Language Acquisition* (henceforth ESK1), investigated the item-based nature of Carlos's *can*-patterns and suggested that formulas, or multi-word expressions (MWEs), should be seen as interactionally and locally contextualized. They were found to be transitory in nature; i.e. their deployment over time was seen to be occasioned by specific usage events. Such events, it was argued, must be recurrent in order for the MWEs, at least the ones identified for Carlos's *can*-pattern development, to be retained by the learner over time. It was further observed that productivity enhancement is partially concrete, based on utterance schema development, and traceable to previous experience. It therefore lies at the heart of the findings in ESK1 that development is not only item-based, but also very much usage-based. It was displayed how patterns wax and wane in response to changing environmental and interactional factors. It follows from these insights that ontogenetic language development is inextricably coupled with language use. A full theory or model of SLA, it was therefore argued, must incorporate room for studying these local contexts in a more detailed manner to investigate in depth the interplay between local interactional contingencies and portable linguistic experience of the individual.

The results in ESK1, then, yielded the research questions to be tackled in the subsequent research papers. The third research paper, *What's new? – Routines and Creativity along a Usage-Based Path of Second Language Learning* (henceforth ESK2), attempted to delineate the characteristics of those aspects of language knowledge that are situated and transitory and those that are durable and portable. In ESK2, the object of research in longitudinal L2 studies was operationalized as a hybrid between locally applied usage patterns and application of the same and related usage patterns over time, and represents the germ of the idea of viewing L2 learning in terms of an empirically

grounded, emergent grammar, consisting of units of spontaneously occurring language use. This resulted in a conceptualisation of emergent creativity as building on recycled linguistic matter in the form of MWEs and utterance schemas, i.e., patterns that are more or less lexically specific. It was shown that the L2 inventory seems to be a structured set of such utterance schemas. Development was thus described and analysed as the emergence of new utterance schemas and the combination of such schemas, in an increasing number of ways with an increasing number of schematically sanctioned lexical options and intra-turn schematic operations.

In ESK2 it was also shown how combinability is about putting together chunks rather than lexical items as separate islands. In other words, the lexical items employed are dependent on the patterns known to the language user – and the patterns seem to have been learnt in lexically specific environments as item-based. This insight, it was argued, equals an empirically substantiated justification of ruling out syntactic combinability as the stuff of learning; instead, item-based utterance schemas were posited as the main linguistic material to learn.

ESK2, then, paved the way towards an understanding of the object of research in longitudinal L2 studies as a hybrid between locally applied usage patterns and application of the same and related usage patterns over time; i.e., an empirically grounded, emergent grammar, consisting of units of spontaneously occurring language use. As such, it stands as a corner-stone in the present research. However, the subsequent research papers were actually to a greater extent fuelled by several questions from ESK1 which had remained unapproached in ESK2, namely questions pertaining to the interdependent nature of linguistic interaction and language learning.

The fourth research paper, *Second language learning as participation and acquisition: towards a new SLA eclecticism* (henceforth ESK3), showed how the use of a MWE, *what do you say*, was initially situated in a recurring environment but later expanded to be used in other environments as well. It was argued that the identical deployment of the utterance in comparable sequences over time, an initial routinisation of the MWE, was an example of Valerio having been socialized into the language classroom practice of inviting for help. This, in turn, was seen to hang together with the learning as participation metaphor, whereas Valerio's ability to use the MWE in a new context was argued to be more profitably thought of in terms of the learning as acquisition metaphor. Based on his activities in the social world of the classroom practices, then, he improved his productivity.

As such, it was demonstrated how a full account of L2 learning needs to take into consideration both participation in social interaction and psycholinguistic notions of cognitive portability, linguistic resources, in terms of acquisition.

Furthermore, it was demonstrated how these learning issues fit nicely into a rethinking of the performance-competence distinction along a time dimension of local performance enhancement (routinisation) and general productivity enhancement which transcends the moment. This introduced an elaboration of Larsen-Freeman's conceptualization of learning as that which is carried across contextual boundaries. *What do you say* was found to be carried across contextual boundaries, time-wise, but only so to a certain extent content-wise; i.e., the expression is locally, interactionally contingent, it does not become relevant in any old conversational situation. The utterance schema that emerges from the MWE, namely the more general *do-schema*, while thus a sediment of those interactional contingencies, is much more generally applicable and not, in terms of use, dependent on a narrowly defined conversational setting. The utterance schema, therefore, is carried across both content-defined and time-defined contextual boundaries. This, in essence, marks the difference, fluid as it is, between 'performance enhancement' and 'productivity enhancement' as I operationalised them in ESK3. Looking back, this distinction between a time-defined and a content-defined notion of pan-contextual portability also seems to apply to the findings in ESK1, where the situated MWEs all seemed to be coupled with certain interactional requirements; i.e., they displayed time-wise portability did not seem to be portable into new environments.

The fifth and final research paper, *You no oich – user-based L2 learning: the case of negation* (henceforth ESK4), further supported the mutually reflexive relationship between linguistic development and interactional requirements. It was shown that Valerio, in usage events requiring him to assist his fellow classmates in getting a task right, used the linguistic resources readily available to him; namely, a non-native-like, lexically specific pattern *you no write*. It was argued that in terms of the UBL framework and the importance it ascribes to issues of type and token frequencies, these usage events which prompted the locally heavy use of *you no write*, may have laid the foundation for what in Valerio's ontogenesis was to become a seemingly statistical feature of his linguistic inventory, namely the co-occurrence in negation patterns of *you* and *no* at the cost of a more native-like *do-negation* pattern. ESK4, then, yet again showed the futility of keeping interaction and learning apart.

The developmental tendencies found in the data for ESK4, *do-negation* patterns for both my focal students, were found to support some of the findings, and refute others, from research in developmental sequences. More importantly, ESK4 questioned the fundamental starting point for defining those sequences, as the data supported the finding from ESK2 that learning L2 syntax is not a matter of context-independent rule-learning across linguistic patterns; rather, it is a matter of construction-dimensional exemplar-deduced tendencies that may or may not become schematized as abstract linguistic knowledge in ontogenesis. It is possible to track pattern development in great detail from the concrete item-based starting point of the patterns to the possible abstraction of regularities that link these patterns as schemas. Such possible abstraction, however, should not be the default starting point for longitudinal L2 learning studies, because, as ESK1 showed, not all patterns lend themselves easily to abstraction.

On a more epistemological note, the application of UBL and its insistence on real usage in real usage events as basis for research has resulted in an empirical substantiation of the futility of keeping 'learning' and 'use' apart in SLA studies. It simply is, as Firth and Wagner (1998) told us, impossible to tell where one ends and the other begins. L2 learning is not as simple as being a matter of restructuring cognitive machinery against the influence of attended input in modified interaction. The model assuming these basics for L2 learners and L2 researchers like, the Interaction/Input/Output-model, has been shown to be inadequate for dealing with the complexities of learners constructing their L2 inventories. Rather, what the data suggest is that linguistic development and interactional requirements are interwoven at a fundamental level; what learners learn is not the outcome of what they do, it is something that emerges in the flux of doing. Learning and doing are mutually reflexive; they happen simultaneously. Some cases in point are the emergence of *you don't verb* in ESK4, the gradual spreading of the *can*-pattern in ESK1, as well as some of the emergent patterns documented in ESK2 (*in + location* and *it's more X*).

ESK3 and ESK4 also showed the need for an elaborate investigative framework to tackle the interactional phenomena encountered in order to capture the essence of the interplay between interaction and learning. For this purpose, the notion of usage event as it is found in UBL did not seem sufficient. Therefore, I invoked Conversation Analysis (CA), or at least some micro-analytic tools inspired by CA, to account for local interactional contingencies found in the data. This

resulted in an eclectic approach to the study of developmental issues in L2 learning, suggesting the need to abolish or rethink certain dualisms; apart from the learning-use dichotomy, these included the performance-competence distinction and the dichotomous relationship between the two learning metaphors, 'participation' and 'acquisition'. It was argued that developmental trajectories found in the data should be analysed and described with recourse to both metaphors, and it was argued, and put into further perspective in chapter 8 on SLA, that an eclectic approach, which attempts to bridge the gaps represented by existing dualisms, is more apt at accounting for L2 learning than one which chooses either side of the dualism fence.

Summing up, it has been argued that UBL is a fruitful framework for exploring developmental issues in L2 learning. The five research papers have, each from different perspectives and with different research objectives, substantiated the UBL claim that language learning is item- and usage-based. The usage-based perspective, increasingly pivotal in the chronology of the five research papers, has called for a non-distinction between interaction and learning. Rather, these have been argued to be fundamentally coupled and happen simultaneously. The item-based nature of the projected trajectory of language learning has implied a view of 'formulaic language' which is different from that found in traditional psycholinguistic research (e.g., Wray 2002; Schmitt and Carter 2004) which has been argued to subscribe to a view of language knowledge as essentially compartmentalized, entailing a definition of formulaicity as something not generated by syntax. UBL, on the other hand, does not distinguish the phenomenon as essentially different from the rest of the linguistic inventory. Language knowledge is fundamentally holistic, all patterns of language uniformly represented in the inventory (Croft and Cruse 2004). This view of language knowledge has been empirically substantiated in L1 acquisition research; Tomasello (2003) put it nicely when he said for children constructing their first language that they must learn two faces of grammar: smaller elements and larger patterns. Now it seems that it is time for the field of SLA to apply this insight and investigate its validity in ever more detail for research on adults accumulating their linguistic resources as they construct their L2 inventories. SLA must rid itself of the compartmentalized view of language; i.e., the idea that lexis and grammar are to be kept apart, in research as well as in teaching. L2 learners simply do not learn the two in a manner that justifies keeping them apart; they are intertwined to the brink of being inseparable, and learners acquire them together, not each in its own paradigmatic vacuum.

## DANSK RESUMÉ

Forskningen som præsenteres i denne samling af fem forskningsartikler og tre baggrundskapitler, blev oprindeligt formuleret med udgangspunkt i et ønske om at udforske fasttømrede sproglige helheder ('multi-word expressions') og deres rolle for L2 læring fra et funktionel-kognitivt lingvistisk synspunkt. På baggrund af empirisk evidens blev forskningen imidlertid hurtigt centreret omkring et ønske om at undersøge frugtbarheden i at applicere 'Usage-Based Linguistics' (UBL) på longitudinel SLA forskning. Grunden til denne ændring var, at spørgsmålet om hvorvidt learners starter fra helheder og gradvist analyserer dem og burger konstituenterne i andre sammenhænge, eller starter fra at lære kombinatoriske regler som med tiden via praksis resulterer i enkelte fasttømrede helheder, syntes forfejlet. I stedet viste mine data, at fasttømrede helheder og generelle mønstre koeksisterer på alle tidspunkter i den ontogenetiske udvikling, i hvert fald så vidt angår mine data og informanter. I det følgende vil jeg opridse resultaterne fra forskningsartiklerne.

Mine data kommer fra 'the Multimedia Adult English Learner Corpus', som består af audiovisuelle optagelser af klasseværelsesinteraktion i et engelsk som andetsprog (ESL) klasseværelse i Portland, Oregon. ESL klasseværelserne hvor optagelserne fandt sted, blev udstyret med seks videokameraer monteret i loftet. Fire af disse var stationære, mens to var fjernstyrede. De to sidstnævnte fulgte hver én kursist som var udstyret med mikrofon; kursisterne skiftedes til at have mikrofon på. Læreren havde altid mikrofon på. Den egentlige database anvendt i de fem forskningsartikler består af transskriptioner fra ca. 70 undervisningsgange, som hver består af tre lektioner, hvor én af mine informanter, Valerio eller Carlos, enten har mikrofon på eller sidder ved siden af én som har. Både Valerio og Carlos er mexicansk-spansk talende engelsk som L2 learners. Valerios data strækker sig fra juli 2003 til juli 2005, Carlos' fra september 2001 til februar 2005.

Som guidende princip i de fem forskningsartikler har jeg anvendt UBL og dens antagelse at sproglæring foregår via en eksemplarbaseret udviklingssti, som går fra sproglige helheder via partielt skematiserede mønstre til fuldt abstrakte konstruktioner. Denne modus operandi viste sig at være ekstremt brugbar og valid fra begyndelsen. I den første forskningsartikel hvor denne fremgangsmåde blev afprøvet, *Are recurring multi-word expressions really syntactic freezes? Second Language Acquisition from the perspective of Usage-Based Linguistics*. (samforfattet med T. Cadierno; EC), fandt vi positiv evidens for UBL's læringssti, hvor systemet der opstår i læringen, er karakteriseret som den gradvise abstrahering af regulariteter, der forbinder konkrete udtryk som

skematiserede konstruktioner. Læringen af *Do-negation* sås som værende meget afhængig af én specifik instantiering af mønstret, nemlig *I don't know*, mens produktiviteten sås som værende stigende i takt med at den underliggende sproglige viden syntes at blive gradvist mere skematiseret, som reflekteret i type og token frekvenser. Disse resultater tydede på, at L2 læring er eksemplarbaseret, at forankring af enkelte faste udtryk er afhængig af token-frekvenser, og at læring af mere abstrakte mønstre og konstruktioner er afhængig af type-frekvenser, som forventet.

Kronologisk set resulterede de enkelte forskningsartikler i indsigter som genererede nye spørgsmål. Den anden forskningsartikel i rækken, *Constructing another Language – Usage-Based Linguistics in Second Language Acquisition* (ESK1), undersøgte om Carlos' læring af *can*-mønstre var eksemplar-baseret og pegede i retning af at de sproglige helheder, operationaliseret som gentagne flerordsforbindelser, skulle ses som interaktionelt og lokalt kontekstualiserede. De sås som værende essentielt transitoriske; dvs. deres brug var afhængig af særlige brugsbegivenheder ('usage events'). Sådanne specifikke brugsbegivenheder skulle også være gentagne for at flerordsforbindelserne, i hvert fald dem som jeg kunne identificere for Carlos' læring af *can*-mønstre, forblev i brug hos learneren over tid. Det blev også observeret at forbedring af den sproglige produktivitet er partielt konkret, baseret på ytringsskemaer (dvs. halvt skematiserede, halvt leksikalsk specifikke mønstre, som f.eks. *I can Verb*) og altid sporbar til tidligere erfaringer. Det er derfor et kardinalpunkt i resultaterne fra ESK1, at L2-læring ikke kun er eksemplar-baseret, men også i høj grad brugsbaseret. Det blev vist hvordan sproglige mønstre tager af og tager til i brug, afhængig af de interaktionelle faktorer i konteksten. Det følger af disse indsigter at ontogenetisk sproglig udvikling er uløseligt forbundet med sprogbrug. En fuld SLA-teori eller –model må derfor inkorporere rum for at undersøge disse lokale kontekster mere detaljeret for at kunne undersøge i dybden samspillet mellem lokale interaktionelle kontingenser og transportabel sproglig erfaring hos individet.

Resultaterne fra ESK1 genererede således forskningsspørgsmålene for de kommende forskningsartikler. Den tredje forskningsartikel, *What's new? – Routines and Creativity along a Usage-Based Path of Second Language Learning* (ESK2), forsøgte at udspecificere kendetegnene for de aspekter af sproglig viden som er situeret og transitoriske, og dem som er langvarige og transportable. I ESK2 var forskningsobjektet i longitudinelle L2 studier operationaliseret som en hybrid mellem lokalt anvendte lingvistiske brugsmønstre og anvendelse af de samme og relaterede brugsmønstre over tid. Artiklen repræsenterer således kilden til ideen om at se L2 læring gennem en

empirisk grundet, emergent grammatik bestående af enheder af spontant anvendt sprog. Dette resulterede i en anskuelse af emergent kreativitet som byggende på genbrugt sprogligt materiale i form af gentagne flerordsforbindelser og ytringsskemaer; dvs. mønstre som er mere eller mindre leksikalt specifikke. Det blev vist at L2 inventariet synes at være et struktureret sæt af sådanne ytringsskemaer. Udvikling blev således beskrevet og analyseret som fremkomsten af nye ytringsskemaer og kombinationen af sådanne skemaer på et stigende antal måder med et stigende antal skematisk sanktionerede leksikalske muligheder og intra-tur-operationer.

I ESK2 blev det også vist hvordan kombinabilitet handler om at sammensætte større helheder snarere end leksikalske enheder som separate øer. M.a.o., de leksikalske enheder som anvendes er afhængige af de sproglige mønstre som sprogbrugeren mestrer – og disse mønstre synes at være blevet lært i leksikalsk specifikke kontekster som eksemplar-baserede. Denne indsigt, argumenterede jeg, muliggør en empirisk substantieret retfærdiggørelse af at udelukke syntaktisk kombinabilitet som det stof som sproglæring er gjort af; i stedet foreslog jeg eksemplar-baserede ytringsskemaer som det kardinale lingvistiske materiale L2 learnere skal lære.

ESK2 brolagde således vejen mod en forståelse af forskningsobjektet i longitudinelle L2 studier som en hybrid mellem lokalt anvendte brugsmønstre og anvendelse af de samme og relaterede brugsmønstre over tid; en empirisk grundet, emergent grammatik bestående af enheder af spontant anvendt sprog. Som sådan står den som en hjørnestein i forskningen som fremlagt i de fem forskningsartikler. Imidlertid var de efterfølgende forskningsartikler i højere grad sat i gang af spørgsmål som stadig stod ubesvarede tilbage fra ESK1, nemlig spørgsmål som omhandlede den nærmere sammenhæng mellem sproglig interaktion og sproglæring.

Den fjerde forskningsartikel, *Second language learning as participation and acquisition: towards a new SLA eclecticism* (ESK3), viste hvordan brugen af en flerordsforbindelse, *what do you say*, var initialt situeret i gentagne kontekster men senere blev udvidet til at blive anvendt i andre kontekster også. Jeg argumenterede for at den initiale anvendelse af ytringen i komparable sekvenser over tid, en initial rutinisering af flerordsforbindelsen, var et eksempel på at Valerio var blevet socialiseret ind i klasseværelsespraksissen 'at invitere til hjælp'. Denne socialisering sås som hængende sammen med 'læring som participation-metaforen', hvorimod Valerios evne til at bruge denne flerordsforbindelse i en ny kontekst blev set som et udslag af 'læring som tilegnelse-metaforen'.



Baseret på sine aktiviteter i klasseværelsespraksisernes sociale verden forbedrede han sin sproglige produktivitet. Som sådan blev det demonstreret at en fuld fortegnelse over L2 læring nødvendigvis må beskæftige sig med både læring som participation i social interaktion og læring som tilegnelse af sproglige ressourcer, som indebærer kognitiv transportabilitet.

Derudover blev det demonstreret hvordan disse læringsfænomener passer fint ind i en gentænkning performans-kompetence distinktionen langs en tidsdimension af lokal performansforbedring (rutinisering) og produktivitetsforbedring som transcenderer øjeblikket. *What do you say* så ud til at blive transporteret på tværs af kontekstuelle grænser, tidsmæssigt, men kun til en vis grad indholdsmæssigt; dvs. udtrykket er lokalt interaktionelt kontingent, det bliver ikke gjort relevant i en hvilken som helst konversationsituation. Ytringsskemaet som emergerer fra flerordsforbindelsen i den eksemplar-baserede læring, nemlig det mere generelle *do*-skema, således et sediment at interaktionelle kontingenser, er langt mere generelt anvendeligt og ikke, så vidt det gælder brug, afhængig af en snævert defineret konversationskulisse. Ytringsskemaet bliver således båret på tværs af både indholds- og tidsdefinerede kontekstuelle grænser. Dette markerer essentielt forskellen, flydende som den er, mellem 'performance enhancement' og 'productivity enhancement' som jeg operationaliserede dem i ESK3. Retrospektivt synes denne distinktion mellem en tidsdefineret og en indholdsdefineret kontekstuel transportabilitet også at gælde for resultaterne i ESK1, hvor de situerede flerordsforbindelser alle syntes at være parrede med særlige interaktionelle krav; dvs. de udviste tidsrelateret transportabilitet men var ikke transportable til nye kontekster, indholdsmæssigt.

Den femte og sidste forskningsartikel, *You no oich – user-based L2 learning: the case of negation* (ESK4), understøttede det gensidigt reflektive forhold mellem sproglig udvikling og interaktionelle krav. Den viste at Valerio i brugsbegivenheder som krævede at han assisterede sine klassekammerater i den igangværende *task*, brugte de lingvistiske ressourcer han havde til rådighed, nemlig et ikke native-like leksikalsk specifikt mønster, *you no write*. Jeg argumenterede at set fra UBLs perspektiv og den vigtighed UBL tilskriver type og token-frekvenser, kunne det se ud som om at disse brugsbegivenheder, som resulterede i brugen af *you no write*, måske har lagt fundamentet for hvad der i Valerios ontogenese skulle blive et tilsyneladende statistisk træk i hans lingvistiske inventar, nemlig sammenfaldet i negationsmønstre mellem *you* og *no* på bekostning af

et mere native-like *do-negation* mønster. ESK4 viste også at de interaktionelle krav ikke kan adskilles fra læringsforløbet.

Udviklingstendenserne i ESK4, der undersøgte *do-negation* mønstre for begge mine informanter, og dermed delvis replicerede EC, sås som støttende nogle af antagelserne, og afvise andre, fra forskning i de såkaldte udviklingssekvenser i SLA. Mere væsentligt satte ESK4 spørgsmålstegn ved det fundamentale udgangspunkt for definitionen af sekvenserne, idet data støttede den antagelse fra ESK2 at L2 syntakslæring ikke er et spørgsmål om kontekstafhængig regellæring på tværs af lingvistiske mønstre; det er snarere et spørgsmål om konstruktionsdimensionelle eksemplar-deducerede tendenser som måske eller måske ikke bliver skematiseret som abstrakt sproglig viden i ontogenese. Det er muligt at spore mønsterudviklingen i detaljer fra det konkrete eksemplar-baserede udgangspunkt til den mulige abstrahering af regulariteter der forbinder mønstrene som skematikker. Sådant mulig abstraktion skal imidlertid ikke være det automatiske udgangspunkt for longitudinale L2 læringsstudier fordi, som ESK1 viste, ikke alle sproglige mønstre lader sig abstrahere problemfrit.

Epistemologisk har anvendelsen af UBL og dens insisteren på ægte sprogbrug i ægte brugsbegivenheder således resulteret i en empirisk substantiering af det frugtesløse i at holde sprogbrug og sproglæring adskilt. Man kan ikke med vished sige hvornår det ene ophører og det andet begynder, så at sige. L2 læring er ikke så net og simpelt at det kan beskrives som et spørgsmål om et restrukturere kognitivt maskineri under indflydelse af input i modificeret interaktion. Modellen som antager disse basale træk i L2 tilegnelsesprocesserne, for L2 learnere og forskere, nemlig 'Interaction/Input/Output-modellen' har vist sig at være inadækvat som referenceramme for kompleksiteterne i konstruktionen af L2 inventarer. Mine data har i stedet peget i retning af at sproglig udvikling og interaktionelle krav er sammenvævede på et helt fundamentalt niveau; hvad learnere lærer er ikke et resultat af hvad de gør, det er noget der emergerer i aktiviteters flygtighed. Læring og gøren er gensidigt reflektive; de finder sted samtidig. Dette ses eksempelvis i det forhåndenværende materiale i emergensen af *you don't Verb* hos Valeiro i ESK4, den gradvise spredning af *can*-mønstre hos Carlos i ESK1 samt nogle af de emergente mønstre dokumenteret i ESK2 (*in + location* and *it's more X*).

ESK3 og ESK4 viste også behovet for et udbygget referenceramme for at kunne beskrive de interaktionelle fænomener i data og for at kunne indfange essensen af samspillet mellem interaktion og læring. Til dette syntes brugebegivenheden sådan som den beskrives i UBL ikke tilstrækkeligt. Derfor trak jeg på Konversationsanalyse (CA), eller i det mindste mikro-analytiske værktøjer inspireret af CA, for at kunne redegøre for de lokale interaktionelle sammenhænge som jeg fandt i data. Dette resulterede i en eklektisk tilgang til studiet af udviklingsmæssige aspekter af L2 læring, som pegede imod nødvendigheden af at gentænke eller forlade visse dualismer; bortset fra læring-brug dikotomien inkluderede disse performans-kompetence distinktionen og det modsætningsfyldte forhold mellem de to læringsmetaforer 'participation' and 'tilegnelse'. Jeg argumenterede for at de udviklingstendenser som mine data viste, skulle analyseres og beskrives i lyset af begge metaforer. Jeg argumenterede ligeledes, og dette blev videre perspektiveret i kapitel 8 om SLA, at en eklektiske tilgang, som søger at bygge bro over de kløfter som dualismerne repræsenterer, er bedre til at indfange og analysere L2 læring end en som vælger den eller den anden side af dualismerne.

Til opsummering: jeg har argumenteret for at UBL er en frugtbar ramme for længdeundersøgelser af L2 læring. De fem forskningsartikler har, hver fra sit perspektiv og med forskellige objekter for øje, substantieret UBLs påstand at sproglæring er eksemplar- og brugsbaseret. Det brugsbaserede perspektiv, i stigende grad vigtigt i kronologien i de fem artikler, har kaldt på en nondistinktion mellem interaktion og læring. I stedet har jeg argumenteret for at disse er fundamentalt parrede og sker samtidig. Den eksemplarbaserede læringssti har impliceret et syn på sproglige helheder som er forskelligt fra det som findes i traditionel psykolingvistisk forskning, som abonnerer på et essentielt kompartmentaliseret syn på sproget, hvilket medfører en definition af 'helhed' som noget der ikke er genereret af syntaksen. UBL på den anden side skelner ikke a priori psykolingvistisk lagrings- og processeringsmæssigt mellem fænomenet helhed og andre typer af enheder i det lingvistiske inventar. Sproglig viden er essentielt holistisk, alle sproglige mønstre uniformt repræsenteret i inventariet. Dette syn på sproglig viden er tidligere, eksempelvis hos Tomasello, i børnesprogsforskningen blevet empirisk substantieret med den overbevisning til følge at børn skal lære grammatikkens to ansigter: mindre elementer og større mønstre. Nu synes tiden at være inde til at SLA søger at anvende denne indsigt og i detaljer undersøger dens validitet for forskning i voksnes akkumulering af lingvistiske ressourcer idet de konstruerer deres L2 inventarer. SLA må frigøre sig fra det kompartmentaliserede sprogsyn; dvs., idéen om at leksis og syntaks skal være skilt ad i forskning såvel som i undervisning og læring. L2 learnere lærer simpelthen ikke de to på

en måde som retfærdiggøre at man holder dem adskilt; de er forbundne i en grad så de faktisk er uadskillelige, og learnere lærer dem sammen, ikke hver for sig i paradigmatisk tomrum.