2
The development of English as a second language

Bruno Di Biase*, Satomi Kawaguchi*
and Yumiko Yamaguchi**
*University of Western Sydney, **Tokai University

1. Introduction

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

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

EUROSLA MONOGRAPHS SERIES 3
Grammatical development in second languages, 85-116

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

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

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

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

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

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

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

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

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

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

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

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

2. Morphological development

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

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

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

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

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

(3a) Kumi's morphological development

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

in brackets time of session in weeks after arrival

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

(3b) Scalability matrix of Kumi's morphological data[2]

		T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14
S-PROCED.	3rd pers sg *–s*	–	–		–	–	–	–	–	+	+	+	+	+	+
NP PROCED.	quantifier + pl *–s*		–	–		⊖		+	+	+	+	+	+	+	+
VP PROCED.	*can't/can/will/couldn't* V		–	–	+			+		+	+	+	+	+	+
	be V*–ing*	–	–	–	+	+	+	+	+	+	+	+	+	+	+
CATEGORY PROCED.	past *–ed*		⊖	⊖	⊖	⊖	⊖	⊖	⊖	+	+	+	+	+	+
	plural *–s*	–	⊖	⊖	+	+	+	+	+	+	+	+	+	+	+
	possessive *'s*		+			+	+	+			+	+		+	+
	V*–ing*	–	+	+	+		+	+		+	+	+	+	+	+

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

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

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

So, what unites the four morphemes within the categorial stage, despite their differences? And in what way are they different from morphemes hypothesised at higher stages? The PT answer[3] lies in whether or not they require grammatical unification with other lemmas in the phrase or sentence structure (Pienemann 1998: 97; and (12), § 2.1, ch. 1, this volume). The four morphemes at the English categorial stage do not require grammatical unification with other constituents of the phrase or sentence in which they may appear, so they belong to the categorial stage. We will now proceed to discuss them following their order of emergence as shown in (2). This order is not a necessary part of the PT hypothesis but a matter of observation in Kumi's data. However, we do attempt to explain, in a principled way, early appearance of some morphemes in this learner, such as *–ing* and possessive *'s*, as well as the late appearance of other morphemes, such as *–ed*.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

The next marker to emerge at the categorial stage is plural *–s* for Ns. For English this means that learners must be able to differentiate first whether the referent is considered as one entity or more (*cat* vs *cats*), and secondly whether the entity referred to is semantically countable or not (*apples* vs *milk*). Learners further need to learn that English requires this *–s* ending form also for referring to generic countable entities (*I like apples*), but not with generic uncountables (*I love coffee*) – not as simple as the learner might hope for, after all. This is especially so for Kumi, whose L1 does not mark plural morphologically. Kumi's data shows these difficulties. The plural *–s* marker emerges at t4, one month later than *–ing* and possessive *'s*, and a pattern of nonsuppliance persists up to t14[6] as the table in (3a) above shows. In terms of structures belonging to the category procedure stage, this indicates the existence of a soft barrier. We note further that t4 is also the time when the phrasal procedure begins to be activated.

---

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

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

The next emerging structure, the so-called regular past *–ed* marker, provides a stronger case of a soft barrier, in so far as it shows up rather late in Kumi's corpus (at t9, about 18 months after arrival in Australia). This is rather late compared to 'irregular' past forms such as *came* and *went* used at t4, which are learned item by item. Once learned, the *–ed* marker is generalised to Vs that do not require it, as in (7a), and to newly learned Vs, whether regular or irregular as in (7b). Then at t14 (the100th week of exposure) the child seems to have the rule under control. This developmental path is common for all learners of English L1, L2, and bilingual L1.

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

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

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

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

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

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

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

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

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

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

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

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

(11) a. *the two black cats*
b. lexical entries

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

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

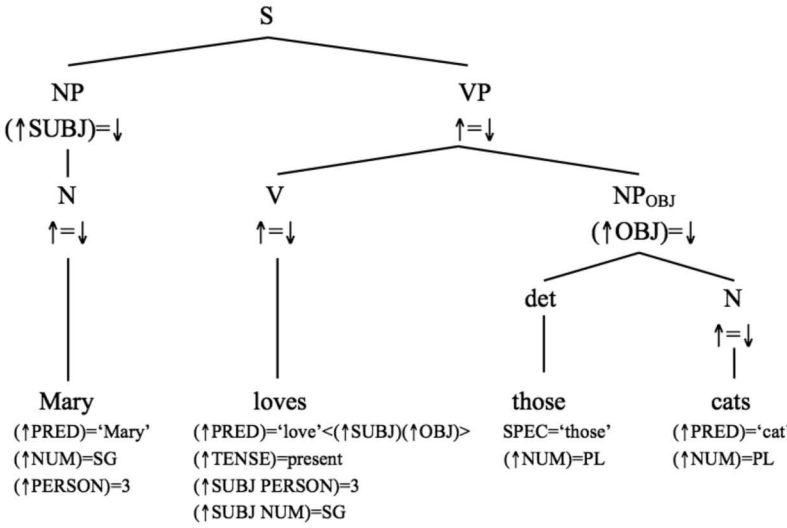
(12) t9 most kids like it

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

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

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

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



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

(14) t9 all the time he wakes up

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

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

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

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

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

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

3. Syntactic development

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

3.1. The Prominence Hypothesis

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

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

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

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

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

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

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

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

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

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

The next step forward in the development of syntax accounted for by the Prominence Hypothesis is achieved in English when the learner's lexicon and categorial differentiation are strengthened and expanded, and TOP is placed in first position. This TOP constituent is not typically a core element of the sentence – that is, it can be a time or place circumstantial ADJ in declarative sentences, to which prominence is assigned by the speaker. This structure comes about early and is target-like[8] in a highly configurational language such as English, where canonical word order is undisrupted in declarative sentences that are pragmatically rea-

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

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

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

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

(19) t13 tomato thing I like it

(20) $S' \rightarrow \underset{(\uparrow TOP) = \downarrow}{(XP)} \quad \underset{\uparrow = \downarrow}{S}$

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

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

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

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

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

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

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

(24) a. did Mary buy a CAR?
b. she bought a motorbike

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

3.2. The Lexical Mapping Hypothesis

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

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

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

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

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

STRUCTURE	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14
unaccusatives														1
passives											(1)		1	
agent/experiencer on SUBJ, goal on O_1, and theme on O_2						1			1	1				
agent/exper. on SUBJ, patient/theme on OBJ, and other arguments on OBL						12	4	2	1	1				
agent/experiencer on SUBJ	6	12	6	4	9	5	8	8	12	3	10	6	5	4
agent/experiencer on SUBJ, patient/theme on OBJ	13	15	31	38	14	15	9	14	28	15	19	10	20	18

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

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

Not before t6 (after about four months' exposure) does Kumi reach the intermediate stage of 'default mapping with an additional argument' where the added argument is marked as such, that is, distinguishing it from SUBJ and OBJ, which are marked by position only. At t6, this additional element is marked both as OBL by a preposition and as OBJ_2 by position, as in (43a) and (43b) respectively.

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

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

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

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

4. Conclusion

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

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

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

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