TARGET ARTICLE

Can the Usage-Based Approach to Language Development be Applied to Analysis of Developmental Stuttering?

C. Savage* and E. Lieven

*Department of Psychology, University College London, Gower St., London WC1E 6BT, England
c.savage@ucl.ac.uk

Max Planck Institute for Evolutionary Anthropology, Leipzig
lieven@eva.mpg.de

Abstract. The usage-based approach to language development suggests that children initially build up their language through very concrete constructions based around individual words or frames on the basis of the speech they hear and use. These constructions gradually become more general and more abstract during the third and fourth year of life. We outline this approach and suggest that it may be applied to problems of fluency control in early child language development.

Keywords: Usage based approach, generativist approach, EXPLAN.

1. Setting the scene

A child begins to control the articulators in an effort to make vocal sounds shortly after birth. These early efforts provide the basis for most language development which develops over the course of three years from single words to simple, and subsequently more complex, grammatical utterances. Any difficulties that a child experiences as she tackles the acquisition of each component skill during language development may result in different types of fluency problem. During the vocabulary spurt at around 18 months, for instance, a coincident increase in naming errors also occurs. This is characterized by perseveration of recently produced words. As well as reflecting the vulnerability of newly acquired items to retrieval error (Gershkoff-Stowe, 2002), it has also been suggested that the errors could reflect the early system-wide fragility of the word retrieval process (e.g. Marchman & Bates, 1994; Plunkett & Marchman, 1993). Over time, with increased practice, these naming errors disappear, suggesting that the demands on the language production system become better aligned with its capacity.

Other problems that occur during the development of fluency control may be more persistent and it is essential to know why. Here we will be examining, from a usage-based approach, what problems in early language development could lie behind stuttering. Stuttering is a disorder that is particularly prevalent in childhood; its modal onset age is 3 years (past the one-word stage of language development studied by Gershkoff-Stowe, 2002). It is possible that the fluency problems experienced at 3 years are a ‘side effect’ of the child acquiring another language skill around this age, similar to the disfluency that results during rapid lexical development. One possible explanation is that grammar is starting to become more general and abstract. Increased disfluency could result from the child’s more sophisticated attempts to combine words into grammatical forms (Bernstein-Ratner, 1997). From this perspective, the high rate of spontaneous recovery during childhood (Andrews, Craig, Feyer, Hoddinott, Howie and Neilson, 1983) could be due to ‘realignment’ of language demands with the child’s capacity to meet them during later language development. The hypothesis is, then, that stuttering prevails at points of vulnerability in language development when the system is under strain through an advance in acquisition of a linguistic skill. However, care must be taken to specify what is meant by ‘grammatical development’ as the particular definition will have ramifications for when this skill is thought to onset. Children combine words from as early as 2 years. They are able to produce novel utterances on a principled but limited basis by, for example, slotting certain nouns or verbs into low-scope frames. This almost always results in utterances in canonical word order, albeit often with material missing (e.g. Lieven, Behrens, Speares and Tomasello, 2003). The nature of children’s early grammar will be discussed in more detail later in this article. The important issue is that, given that children gradually develop their ability to combine words grammatically from before 3 years, why do the problems in speech control start relatively late (i.e. at 3, rather than 2, years)? The current article will provide a developmental perspective on stuttering and speech production in general. The nature of children’s early linguistic knowledge will be discussed (section 2), and how this knowledge relates to online speech production (section 3), the mechanisms of speech production and how disfluencies could arise are
discussed in section 4. A specific theory about the development of stuttering (EXPLAN) is reviewed in section 5, and how this theory might relate to the language development models reviewed in section 2 (section 6). Finally future directions of research are proposed that would take into account the findings from all these fields in a unified way (section 7).

2. The Nature of Early Linguistic Representations and its Relationship with Language Production

The modal onset age of stuttering is 3 years yet syntactic development precedes this and has progressed beyond the one word stage. The reason that age of stuttering onset does not coincide with age of language onset may have important ramifications for the language acquisition debate. Stuttering begins during early language development, suggesting that an adequate explanation of stuttering must include a developmental perspective. The critical question here is how to characterise children’s underlying linguistic competence across language development and how to measure it. Studying the abilities of young children places restrictions on the techniques we can use. As always in developmental research, much of the challenge for researchers lies in finding appropriate methodologies. Standard tests of language development are not widely available for children of this age and the usefulness of a gross measure such as mean length of utterance (MLU) is limited. MLU is often measured on naturalistic data but a single MLU measure on a small sample of speech is not sufficient to characterize a child’s state of language development. MLU does not differentiate sufficiently between fine-grained distinctions of language sophistication. So, for example a high MLU score might be obtained for a child who strings together formulaic and phonologically simple units using “and” (example 1 below) compared to a child who produces shorter phrases but does so by using more abstract and creative production schemas (example 2). Thus the creativity and the length of an utterance do not necessarily concur.

1. Mum laughed and Dad laughed and Bob laughed (11 morphemes).
2. I bloomed these flowers (= made these flowers bloom) (6 morphemes).

These examples indicate that MLU is not ideal for examining factors that may affect fluency, such as syntax and phonological complexity. In this paper some of the methods used to study children’s naturalistic language development will be explored in depth, with particular reference to the ‘Usage-based Approach’ (Langacker, 1987) (UB approach). We will examine the relationship between the existing work on fluent language development with that on developmental stuttering and we will suggest how the UB approach might provide new methods for studying stuttering near onset. First, however, we outline the potentially very different predictions about disfluency that stem from the UB approach and the traditional formal, approach to language development.

The generativist approach (starting with Chomsky’s work in the 1960s) conceptualises linguistic competence as based on Universal Grammar (UG), and maintains that children have the abstract categories of UG from the start. In contrast, the UB approach suggests that children initially lack abstract knowledge and instead start their grammatical development by building up lexically-specific patterns. The basic problem with the generativist theories is that they analyse the child’s language in terms of the categories of adult grammar, thus ‘defining away the problem’. There are, however, a large number of studies that suggest that children’s early language is precisely not that of the adult grammar (see sections 2.1. and 2.2. below). Interestingly, the two assumptions make different predictions for both the age of stuttering onset and the location of disfluencies in speech. The generativist approach predicts: (a) Children would stutter as soon as multiword utterances occur if the problem stems from syntax; (b) Children would stutter around phrase structure boundaries (e.g. before noun phrase, NPs, or verb phrases, VPs or prepositional phrases, PPs) because these are the points in utterances where speech is planned. The predictions of the UB approach are different: (a) Stuttering may start at a later age (around 3 years) because experimental evidence (see section 2.2.) suggests that children only start to be productive with the more abstract aspects of grammar e.g. the argument structure of verbs, towards the end of their third year; (b) If children stutter they will do so at the boundaries of psychologically real planning units that may not always coincide with those of traditional syntactic theory. Thus UB theory maintains that frequently used patterns will continue to be represented lexically into adulthood as well as being analysable in terms of abstract syntax (e.g. “I dunno” and “I don’t know”, Bybee and Scheibman, 1999). The implication of this is that stuttering could occur within a constituent if it exists as a representational unit (e.g. I wanna) as the child attempts to combine previously learned frames (I wanna with biscuit leading to “I wanna bbbbibiscuit” instead of “I want aaaaaa biscuit”, where the stuttering occurs at the NP phrase boundary).

The generativist approach to syntactic development does not predict the modal age of stuttering onset (3 years), whereas the UB approach may provide an explanation for this age of onset. Research within the UB approach suggests that, for normal development, 3 years is the stage in development
when grammar is being reorganised from lexically-specific schemas into more abstract syntax. Stuttering would then result as a ‘side effect’ of the increased demands that this reorganization places on the child’s processing system. Later, the child’s speech production system has adapted to perform these operations more efficiently, and the older child will achieve greater fluency. The proposal is a grammatical parallel of Gershkoff-Stowe’s theory of how fluency is affected by the development of lexical access.

The question remains open as to why some children continue to stutter for longer than others and why a minority continue to do so into adulthood. There are processing theories that have begun to answer this question, which will be explored later. The UB approach may provide an original avenue for exploring differences between children who stutter and fluent speakers across development. For instance, are there differences in the size of the units that children who stutter store and employ relative to fluent children? Are children more likely to speak fluently if they use more concrete and directly accessed language units? How does the development over time of more abstract units relate to creativity and efficiency in language production? The same methods that have been applied to ascertain the nature of early syntactic knowledge in normal language development could be used to help answer similar questions for children who stutter. A discussion of the existing research for normally-developing children will make more concrete the second prediction made above about the location of disfluencies by revealing the planning units that young children use.

Observational studies of naturalistic child language data in the 1970s indicated that young children use at least some of their language in item-specific ways (e.g. Braine, 1976). Early attempts to explain the data proposed that young children arranged their linguistic knowledge around limited scope cognitive-semantic categories that they learned from their language input. Adult-like verb-object pairs appeared in the data but seemed to belong to a variety of independent positional patterns, such as “see + X”, “want + X”, and “have + X”. Braine suggested that each pattern was associated with a specific semantic content, for example a pattern related to oral consumption, associated with eat, bite, and drink. The early cognitive-semantic theories were the first attempts to radically rethink the fundamentals of the language acquisition problem after the generativist tradition emerged. However, subsequent research revealed that children’s use of grammatical structures crosses the boundaries suggested by semantic groupings (e.g. Levy, 1983; Valian, 1986; Pine, Lieven and Rowland, 1998). The UB approach has subsequently pioneered a more thorough approach to data that allows more accurate identification of the ways in which children’s early grammatical knowledge is restricted. The approach adopts rigorous analytic methodologies that use consistent statistical criteria to classify structures within naturalistic data and backs these up with experimental data. Without this systematic treatment of the data there is a risk of either under-, or over-, estimating what a child can do (e.g. experiments that are too hard for the child to perform or naturalistic analyses that accept a structure as acquired at an abstract level on the basis of one exemplar).

2.1. Naturalistic Data

In working with naturalistic data there is always the problem of how to determine whether a structure is productive or rote-learned. Of course, without a 100% sample of what the child says and hears, this can never be done with full certainty for any utterance. But this is all the more reason to develop consistent definitions and to apply them uniformly. However the definition of how to classify a structure as having been acquired has not been consistent either within or between analyses. The number and type of instances that are observed in a corpus are interpreted differently depending on the theoretical framework within which an author is working. Radford (1990, 1995, 1996), for example, draws a distinction between ‘acquisition’ and ‘mastery’. After acquisition but before mastery, children alternate between correct and incorrect use, and sometimes omit functors whilst learning exactly how to manipulate their newly acquired grammatical knowledge (Radford, 1990, 1995, 1996). Gathercole and Williams (1994) point out, though, that Radford applies the criteria non-uniformly. The same type of utterance is classified differently depending on which stage it appears in. For example he describes a ‘How are you?’ that occurs during his ‘lexical stage’ as rote learned because it is repeated monotonously in the same transcript and is the only example of a correct wh-question. In the same ‘lexical stage’, conversely, he classifies as productive sporadically produced questions such as ‘Doing what?’, which could just as well be rote learned from adult echo questions like ‘You’re doing what?’.

Furthermore, his functional stage evidence for adjectival complements rests on only one example. A different example comes from Valian (1991) who based category assignment on the surrounding linguistic and social context of an utterance, and assigned a word in a child’s utterance to the same category that it would be assigned to in adult speech. The basis for this judgment was adult grammar, which effectively rules out other possible interpretations. Valian’s work assumed a priori the existence of the categories for which evidence is being sought, making the argument circular.
The UB approach tackles the problem of non-uniformity in data analysis by abandoning the traditional distinction between underlying competence and surface performance. The utterances of the child are instead treated as a direct indication of the child’s underlying representations. The justification is that a canonical utterance produced by a child cannot indicate how abstract her underlying knowledge is. She may have produced it on the basis of an abstract generative grammar or she could just as easily have produced it directly from a string of words and phrases previously heard and learned. There is no absolute method of distinguishing between the two accounts in the case of naturalistic data because both routes would produce an adult-like utterance (either by knowing adult-like grammar or by hearing adult-like speech in the input).

The only evidence in naturalistic data from which we can conclusively infer abstract knowledge is the creative language use seen in examples of overgeneralization errors like the one given in (2) above and ‘She giggled me’ (Bowerman, 1982). These have been systematically studied and reported for only a few children and they only occur with any frequency after 3 years of age with almost none before 2.5 years (Pinker, 1989). So what is the nature of children’s representations before 3 years? After the early work on cognitive-semantic, item-based patterns in early language, more recent models have suggested that early utterances follow input-based lexical patterns (e.g. Braine, 1988; Pinker and Lieven, 1993; Lieven, Pine and Baldwin, 1997; Pine, Lieven and Rowland, 1998; Tomasello, 1992; Tomasello, 2000). One of the first influential accounts in this tradition was Tomasello’s (1992) ‘Verb-Island Hypothesis’ (VIH). This was based on a diary study of his English-speaking child, T, between the ages of 15 and 24 months. T’s early language use was conservative, with little overlap in the way she used different verbs but much continuity of use within a single verb. There was no evidence of a verb general category. Some verbs were never marked for tense (past tense morphology in English) or aspect (the use of auxiliaries and morphology to indicate the nature of a situation, e.g. whether it is complete or ongoing, fixed or changing, temporary or of long duration), others were marked only for one or only for the other, and very few (2%) were marked for both functions. An example is that T first used *spill* in its past tense form only but *drop* in the present-tense only, before they later became used in both ways). Tomasello’s (1992) VIH was that children learn language conservatively and initially acquire individual verbs, tied to the structure in which they appear in the input. Very young children (2 years) lack any abstract knowledge of either verb categories or constructions. Their language use is built around verb-specific schemas with open nominal slots, for example young children would represent the transitive verb ‘hit’ as:

\[
\text{Noun Phrase (hitter)} \rightarrow \text{form of verb HIT} \rightarrow \text{Noun Phrase (hittee)}
\]

The adult-like verb-general entities ‘subject’, ‘object’ and/or ‘agent’ and/or ‘patient’ are replaced by the verb specific ‘hitter’ and ‘hittee’, ‘pusher’ and ‘pushee’ and so on. Children acquire individual verb meanings incrementally and only later do they begin to form abstractions by generalizing across verb forms.

Lieven, Pine and colleagues (e.g. Pine and Lieven, 1993; Pine, Lieven and Rowland, 1998; Theakston, Lieven, Pine and Rowland, 2001) later showed that the VIH’s focus on verbs was too narrow. Pine, Lieven and Rowland (1998), for example, investigated the extent of lexical overlap shared between different instances of grammatical construction types in the first 6 months of twelve English-speaking children’s multiword speech. They found that although most of the children’s language was grammatical, much could be explained by a relatively small amount of lexically specific knowledge. On the one hand, Tomasello’s research was supported in that the children exhibited a lack of overlap in the verb types to which they applied different morphological markers, lack of overlap in the verb types with which they used different auxiliaries, dominant usage of the first person singular nominative pronoun (I), and lack of overlap in the lexical items used as subjects and direct objects of transitive verbs. On the other, children did show some grouping of verbs into relatively narrow subgroupings based on low-scope frames with slots, for instance “I’m VERB-ing it”. However, this lexically-specific use of grammatical items provides no evidence of truly abstract syntactic categories or schemas. Rather, the data suggest that young children’s knowledge is non-general and organized not only around verbs but also around other high-frequency markers, especially pronouns.

More recently Lieven, Behrens, Speares and Tomasello (2003) were able to explore in further detail the form children’s early lexically organised schemas may take by analysing a dense dataset from a single 2-year-old fluent girl, Annie. They assessed the relation between one sample of Annie’s creative utterances (one hour of interaction with her mother at age 2;1.11) and those that she had previously produced (the previous 6 weeks of samples at a rate of 5 hourly sessions per week, and the accompanying maternal diary of novel utterances). In the hour-long sample at age 2;1.1.1. there were 295 multi-word utterances of which 37% were ‘novel’ (not previously produced in their entirety). These novel utterances were compared with utterances from the previous six weeks of the same child’s...
data using a ‘morpheme-matching’ method to identify how each novel utterance in the final session differed from its closest match in the preceding corpus. This systematic method was used to identify schemas in Annie’s speech, by considering the number and order of the morphemes that were shared between utterances and the frequency of utterances of a similar form. If a number of utterances was found with the same positive morpheme match and the same overall number of morphemes, but with type variation in the same position as the target utterance, then this was defined as a schema, for example ‘I got the butter’ and ‘I got the door’ (matching morphemes underlined) would constitute instances of a ‘I got the W’ schema (W indicates a word that varies across utterances).

Lieven et al.’s (2003) data suggest that the apparent high degree of creativity exhibited in the naturalistic speech of young English-speaking children could be at least partially based on entrenched schemas and a small number of simple modifying operations. There was a very close relationship between Annie’s novel utterances and her preceding utterances, with most requiring only the substitution or addition of a single word. These findings do not seem to be accounted for by the repetitive nature of child-parent conversation, because the application of the same process to Annie’s mother’s speech revealed that though she produced a number of novel utterances in the last session comparable with Annie’s, the preceding corpus did not provide as many prior schemas and exemplars, and more of the matches required multiple and more complex operations, such as insertions and rearrangements. Lieven et al.’s (2003) analysis can be directly related to children’s on-line language production (see discussion in section 3 below).

In summary research using children’s naturalistic data is highly suggestive of a period in early language development when children are building up low-scope patterns on the basis of their pragmatic understanding, what they want to say and relative frequencies in the input. For sure, children are creative with language from the beginning but the UB approach suggests that the scope of their creativity changes from initially low-scope, item-based representations and slowly develops into the more general and abstract knowledge represented by adult language. However, to establish this experimental research is required that can control what it is the children are hearing and identify how their generalisations are being made.

2.2. Experimental Data

Naturalistic data are needed to provide insight into the language patterns that children actually choose to use in their speech, which is particularly useful for studying which factors affect fluency. They provide language in context rather than artificially isolating words or structures which could miss certain features of the data by narrowing the field of observation too much (e.g. ignoring relevant constructions) or making the stimuli too hard or too easy (e.g. the child who stutters either exhibits no disfluencies or cannot say the words at all). This said, it is crucial to back up naturalistic findings with experimental evidence when the concern is to identify how abstract children’s early representations are. Experimental work is the only way to test predictions and identify causal links. Any conclusions from naturalistic data are correlational and carry the risk of suggesting spurious causal relations created by hidden variables. Several inventive experimental methodologies have been developed. Each reveals different aspects of children’s knowledge and, in addition, different methods work better for different ages. In most paradigms, the data can only indicate how abstract children’s representations are if nonce material is used to systematically control the input that children receive. This allows the language variables that affect performance to be teased out and determine how the language input affects the language output. In contrast, using familiar words is subject to the same problems as naturalistic data (i.e. the words are likely to have been heard by the child already). Only by using novel nonsense (nonce) words (e.g. ‘dacking’) can the investigator control the ways in which the child has already heard the words used and be sure that any productive use (i.e. in a construction different from that which they have heard in the experiment) stems from at least partially abstract generative patterns.

One experimental paradigm used to test comprehension is ‘acting out’. Experiments have shown that from early in development young children show the ability to act out some types of sentence appropriately, for example, the English transitive, but only when they use familiar verbs, not unknown verbs (e.g. Roberts, 1983). When 2-year-olds are taught a novel action paired with a novel verb (e.g. ‘This is dacking’) and are then asked to ‘Make X dack Y’ or ‘Show me: X is dacking Y’ they are equally likely to make either X or Y the agent of the action (Akhtar & Tomasello, 1997).

Acting-out can only be done with children from about 2 years of age. Before this the preferential looking methodology is often used. This uses infants’ looking patterns to infer their ability to distinguish utterances with contrasting syntax. Using designs of this kind, researchers have found that some 2-year-olds seem to be responsive to some aspects of transitive constructions in English if they are used with verbs that they know. For example, Naigles (1990) showed that when children as young as 2;1 heard transitive utterances with known verbs they preferred to look at one participant doing something to another (indicating a causative meaning) rather than two participants carrying out
synchronous independent activities (intransitive meaning). It is important to note, however, that this study does not tell us that 2-year-olds possess a full awareness of the functional role of the verb (for example to be able to connect the pre-verbal position with the subject and the post-verbal position with the object) because the test was conducted with familiar verbs. The children could simply be working on the basis of the patterns in which they had previously heard the verb used rather than working on a verb-general template of the transitive. Only Fisher (2000) has used unknown verbs in the preferential looking paradigm, using sentences like “The duck is gorping the bunny up and down”. However, the sentences she presented to children (1;9 and 2;2) contained prepositional phrases that provided additional information that children could use to interpret the meaning of the sentence instead of using the formal syntactic marking (see also Fisher, 1996, 2002). The precise nature of the knowledge that children use in these preferential looking tasks and how abstract it may or may not be is still under intense investigation.

There is also a large and growing body of evidence on children’s early linguistic knowledge from production experiments using novel verbs. Most of these focus on children aged between 2 to 3 years because of the significant developments in grammar that occur at this time. The overall finding is that 2-year-old children’s early productivity with syntactic constructions is highly limited. This confirmation comes from studies like that of Tomasello and Brooks (1998), who exposed 2- to 3-year-old children to a novel verb used to refer to a highly transitive and novel action in which an agent was doing something to a patient. In the key condition the novel verb was used in an intransitive sentence frame such as “The sock is tamming” (e.g. a bear was doing something that caused a sock to ‘tam’, which was akin to rolling or spinning). Then, with novel characters performing the target action, the adult asked children the question “What is the doggie doing?” (the dog was causing a new character to tam). Agent questions of this type encourage a transitive reply such as “He’s tamming the car”, which would be creative since the child had only heard this verb in an intransitive sentence frame. The results showed that very few children produced a full transitive utterance with the novel verb. As a control, children also heard another novel verb introduced in a transitive sentence frame, and in this case virtually all of them produced a full transitive utterance, demonstrating that they can use novel verbs in the transitive construction when they have heard them used in that way. Moreover, 4- to 5-year-old children are quite good at using novel verbs in transitive utterances creatively, demonstrating that once they have indeed acquired more abstract linguistic skills children are perfectly competent in these tasks (Pinker, Lebax & Frost, 1987; Maratsos, Gudeman, Gerard-Ngo & DeHart, 1987; see Tomasello, 2000, for a review).

In another novel verb experiment with young children (aged 2;5 to 4;5), Akhtar (1999) used the novel verb to describe a transitive action but used it with a novel word order, for example “The bird the bus meeked”. When the younger children were given new toys and encouraged to talk about the new event, they quite often repeated the pattern and said things like “The bear the cow meeked”. On the other hand, the 4-year-olds consistently corrected the pattern they had heard to canonical English word order in their responses (e.g. correcting to “The bear meeked the cow”). These findings are consistent with the hypothesis that when 2- to 3-year-olds learn about meeking they only learn about the order of arguments for meeking and do not assimilate the newly learned verb to a more abstract, verb-general linguistic category or construction that underpins the canonical English transitive. However it is clear that this is a developmental process in which knowledge of the transitive slowly builds up. Thus the 2-year-olds were better at repeating sentences with novel verbs if they had heard these verbs used in SVO order than if they had heard them used in the ungrammatical (for English) orders SOV or VSO. This suggests some sensitivity to conventional usage. Abbot-Smith, Lieven and Tomasello (2001) obtained similar results for younger children with the intransitive construction. When children aged 2;4 were presented with a verb they knew (e.g. jump) in noncanonical word order they corrected the word order they heard to canonical word order even more than the 2-year-olds in Akhtar’s study, but they were more likely to use the ungrammatical word orders that they had heard with novel verbs. However, they too more readily reproduced canonical orders than noncanonical orders. This suggests that they possess some knowledge of an abstract transitive construction but that this is not yet strong enough to use to override the patterns they receive in the input.

Overall, the results from the different methodologies indicate that between 2 and 3 years of age, young English-speaking children are still in the process of building up their abstract, verb-general constructions. Similar evidence also exists for other languages (e.g. Allen, 1996, for Inuktitut; Pizutto and Caselli, 1994, for Italian; Rubino and Pine, 1998, for Brazilian Portuguese). The results so far mainly relate to what children lack before 3 years of age with little implication about the ways in which they are productive with language in their multi-word utterances. The naturalistic data discussed above (Tomasello, 1992; Pine & Lieven, 1993; Pine, Lieven & Rowland, 1998; Theakston, Lieven, Pine & Rowland, 2001; Lieven, Behrens, Speares & Tomasello, 2003) suggest that though 1- and 2-year-olds lack abstract syntactic knowledge in the sense of UG competence, they are able to use a more rudimentary grammatical system of lexically-based productive patterns, like ‘More X’ or ‘I’m Xing it’.
Experimental evidence exists to back this up. Data-driven learning theorists reason that these ‘slot-and-frame’ schemas could be derived from a combination of the repetition and systematic variation of phrases that have been found to occur in the input (Cameron-Faulkner, Lieven & Tomasello 2003). The consistent parts could be imitatively learned whilst the more abstract slots could be formed on the basis of the communicative parallels across varied instances. The $X$ in ‘I’m Xing $it$', for example, refers to a different action in each instance but its communicative function is constrained by that of the whole schema and always refers to acting on an object.

Children and Tomasello (2001) investigated the building of slot-and-frame schemas experimentally. They trained 50 children aged 2;6 with several hundred transitive utterances using either 16 familiar or 16 unfamiliar English verbs (as measured by whether or not they appeared for 2-year-olds on the MacArthur CDI; Fensen, et al., 1994) spread over 4 sessions in as many days. For some children, the agent and patient in all sentences were labelled with only nouns and for other children they were labelled with both nouns and pronouns. The children then saw 4 novel actions and heard 4 novel verbs in non-transitive constructions (2 intransitives and 2 passives, using one N and one PN in each). To see if they could produce a transitive with the novel verb, children saw each novel action modelled again and were asked both neutral and transitive-pulling elicitation questions.

The authors found that, regardless of whether the verbs used during training were familiar or unfamiliar, children were best at generalising the transitive construction to the novel verb if they had heard both pronouns and nouns during training rather than nouns only, for example they generalized better after hearing “Look! The bear’s striking the tree. See? He’s striking it” than after hearing “Look! The dog’s hurling the chair. See? The dog’s hurling the chair.” This suggests that the children either learned a more abstract transitive schema from which to generalise to new verbs by hearing variation in the subject and object slots, or that they learned a pronoun-based transitive schema. One feature that supports the latter conclusion is the fact that after the pronoun training condition, children mostly used pronouns in their responses. This supports the suggestion raised above when discussing naturalistic data that pronouns are important in forming schemas.

Further experimental evidence that children’s syntactic knowledge about transitives is arranged around pronouns comes from a priming paradigm (Savage, Lieven, Theakston and Tomasello, 2003; submitted). Savage and colleagues recently adapted the ‘syntactic priming’ paradigm and used it for the first time with young children. The paradigm is well established in the adult literature (e.g. Bock 1986; Pickering and Branigan, 1999; Schenkein, 1980). It refers to the phenomenon whereby processing an utterance with a particular syntactic structure facilitates the processing of a subsequent utterance with the same or a related syntactic form, even though alternative forms would be semantically appropriate. An example would be when a speaker chooses the passive over the active after producing another passive recently. The consensus view is that that syntactic priming reflects underlying syntactic knowledge and as such it provides a methodology to tap into speakers’ underlying knowledge (e.g. Chang, Dell, Bock and Griffin, 2000; Pickering, Branigan, Cleland and Stewart, 2000).

For adults, a stronger priming effect is produced when the primes and targets share the same verb than when they do not (Pickering & Branigan, 1998). Having said this, the effect is significant even when the verb differs between the primes and targets (Pickering & Branigan, 1998). The latter finding, in particular, demonstrates that priming is effective over and above an effect of lexical overlap and indicates that adults possess abstract knowledge of the primed construction. For children, however, the pattern is different. Young children (3-4 years) only show a priming effect when there is much lexical overlap between the primes and targets, for example by priming them with a slot-and-frame type schema like ‘It got V-ed by it’ for the passive, where they need only slot the target verb into the lexical pattern to be primed. Later on however (by around 6 years), children are primed in a more adult-like fashion. They show an effect both in the high lexical overlap condition and also at a more abstract level, for example ‘The bricks got pushed by the digger’ will prime ‘The cake got cut by the knife’, where both the N and V slots vary between prime and target. These data confirm that young children lack adult-like abstract syntactic knowledge, but interestingly they also indicate that they do possess some verb general knowledge. They are thus in line with the naturalistic data of Pine, Lieven and colleagues that children’s early syntactic knowledge takes the form of slot-and-frame schemas that are anchored by lexical items like pronouns.

To conclude this section, the experimental evidence converges with the naturalistic data in indicating that children’s early syntactic knowledge takes the form of slot-and-frame schemas. Only in the months leading up to 3 years do children start to form relations between these schemas and build more abstract knowledge of constructions. The evidence suggests that children of 3 years are reorganising grammar from lexically specific schemas into more abstract ones. This coincides with the age at which children become disfluent. This supports the suggestion made earlier that stuttering could therefore be a ‘side effect’ of the increased demands that this reorganization places on the child’s processing system. In the following sections, the changing nature of children’s syntactic knowledge across development is
considered and how this may relate to their fluency behaviour in terms of the online speech mechanisms involved.

3. Low-Scope Schemas and how they Relate to Fluency in Speech Production

The relationship between linguistic representations and language production in real time is complex. A growing body of evidence exists that suggests that adults possess more than one route of access to their underlying representations during speech (e.g. Bybee & Scheibman, 1999), meaning that as well as being able to generate novel utterances from abstract representations they also seem to produce ‘frozen’ strings of language more directly without using abstract grammar, so called ‘unanalysed’ units. One example is ‘grammaticalization’, the diachronic process by which frequent usage allows conventionalized structural expressions to emerge in a linguistic community (Bybee & Hopper, 2001). Over time the link between a phrase and its underlying form is lost, for example ‘(be) supposed to’ has lost its passive status. In its reduced form the infinitive is phonologically fused to the verb such that a passive agent seems grammatically unacceptable, as in ‘He’s s’posed to be very knowledgeable *by most people’ (where asterix means grammatically unacceptable) (Bybee & Thompson, 1998, p. 2). The position this issue is given in theories of language production depends on the perspective taken, but its importance is recognized in both UB and generativist theories. Generativist theories retain the central place of abstract representations but recent versions do not simply ignore direct access as being peripheral. Some authors have become increasingly involved with explaining the relationship between generativist grammar and the role of the ‘exceptions’ (idioms, partial idioms and low-scope constructions) (e.g. Lebeaux, 1988; Jackendoff, 1996; Culicover, 1999). However, in UB theories of grammar, the idea of different degrees of abstraction is an integral component, as will become clear below.

The UB approach has its roots in cognitive and functional linguistic theories that view grammar as a response to discourse needs. As a result grammar is dynamic and experience-driven (e.g. Bybee, 1998; Goldberg, 1995, 1999; Hopper, 1987; Langacker, 1987). These authors characterise grammar as an inventory of language specific ‘constructions’, or schemas, for example the ‘get’ passive in English could be ‘NP got V(past tense) by NP’. Some authors would argue that these constructions can be at a highly abstract level that parallels the abstract nature of traditional grammar (though, importantly, they are not identical because they are still language-internal rather than universal) (e.g. Bybee, 1998; Ono and Thompson, 1995). Others maintain that a high level of abstraction is unlikely (Croft, 2001). Ultimately these two positions will have to be resolved empirically. The important point is that UB theories all agree that redundancy exists in underlying representations. That is, the existence of a directly accessible, construction-level representation (multiple words stored as one unit) does not preclude the co-existence of smaller level units (i.e. words) that can be used to generate sentences when a more unusual linguistic composition is required. Language users might be able to coincidentally recognise both construction level units and their individual component parts, and the construction and its units may be interlinked in memory by means of a ‘network’ of representations (Bybee, 1998).

The most unanalysed chunk at the most concrete level of representation for adults would be an idiom, such as ‘He kicked the bucket’ to mean ‘He died’ (and possibly some additional semantic nuances), which cannot be understood by combining the individual words but can be recognized as containing analyzable components that appear elsewhere in similar syntactic positions. Note that while this is an idiom it can be used in interaction with other constructions to produce novel utterances, for instance in different tenses (He has kicked the bucket) or utterance-level constructions (Did he kick the bucket?). Another example is the various levels at which a person could represent the passive. The ‘get’ passive could be represented as a concrete, utterance-level representation without a by-phrase (a set phrase derived from the input, such as ‘It got broken’). It could also occur as a partially abstract item, with some filled (in bold) and some open slots as in ‘It got Verb-ed’. This representation could in turn be connected in memory to that of the by-phrase so it could occur both with and without it (It got verb-ed by X). At the highest levels of abstraction the construction would be represented at a more general level for all component parts and no concrete parts, and it would be connected to the be-passive or even to other constructions (Tomasello, 2003), for example ‘N be/get (any tense) Ved (by N) (He was eaten by a dragon; He got run over by a bus). From the ‘network’ perspective (e.g. Bybee, 1998), grammar exists as an interlinked inventory of constructions ranging from fully concrete, to partially concrete, to fully abstract. The links between concrete phrases represent the building up of levels of abstraction, but the sub-units can also be accessed directly, depending on how creative and original an utterance the speaker wishes or needs to generate. On occasion, it may be that there is nothing directly useful in the existing inventory, or that it cannot be recalled. This might be how errors of commission and overextension arise. The network would function via a sort of ‘path of least resistance’ operation for language production.
When adults or children produce speech in real time (online), a complex interaction takes place between the different routes of access to representations. The relevance of the UB models to fluency research is that they should be able to generate concrete predictions about where disfluency is most likely to occur (i.e. in which parts of a construction and in which forms of construction). Which phrases have to be built up and which can be accessed directly would impinge on fluency. In particular, the more concrete phrases will require less effort from the language production system (making them less prone to disfluency) than the more creative utterances. The more frequently used and more concrete phrases demand only a relatively direct and automated retrieval step, whereas more creative productions require that certain abstract slots are filled with lexical items. The moments in speech that precede the more abstract parts of an utterance will be more vulnerable to disfluency as a consequence of the greater demands placed on the language production system at that point, in preparation for the subsequent part of the utterance. The loci of these vulnerable moments will vary for adults and children, because children’s representations are gradually changing over time.

The findings from Lieven et al. (2003) that were discussed above suggest that in children’s on-line language production, concrete strings stored in memory might interact with more abstract categories that have traditionally been the central focus of linguistic study. The knowledge required to underpin Annie’s language processing could be a combination of strings stored in memory that provide lexically-specific ‘frames’ and categorical knowledge that is used to fill ‘slots’. Lieven et al. (2003) tentatively propose that the types of schemas they identify are psychologically realistic in terms of how a child constructs novel utterances on-line, constituting storage and planning units. Likewise, they propose that the operations they identified may represent psycholinguistic operations that children use to manipulate their representations and construct novel utterances on-line. Examples are ‘Substitute’ (using the same lexical pattern with one item substituted for another, as in the formation of “I have some coke” and ‘Add-on’ (the same lexical pattern plus one added item, as in the formation of “Put a bit more here” from “Put X” and “A bit more here”). Whether these operations exist is an empirical question. One way of exploring it would be to investigate the impact of the operations on speech fluency. This would involve using data on stuttering to inform theories of fluency development. Conversely, though the schemas and operations identified by Lieven et al. (2003) were found for a fluent speaker, their analysis could provide an interesting new angle on how to explain fluency and disfluency patterns in speech production also for children who stutter. A sophisticated and empirically-based model of speech production in children could be developed by considering how the schematic operations suggested by Lieven and colleagues interact with the child’s developing mastery of phonetic complexity (e.g. MacNeilage & Davis, 1990; Davis & MacNeilage, 1995; Jakielki, 1998) and phonological representations.

The above proposals are as yet unsubstantiated. The empirical evidence to date for UB models has focused mainly on the nature of the underlying linguistic representations that children and adults possess and not how they use these to produce speech. It is still an open question as to whether the models of representation have any real psychological significance in online speech production. There are, however, a few studies that relate the UB approach to online processing in adults. Schilperoord and Verhagen (1998) studied the timing of speech production when adults read a passage of text. They showed that the locations at which their subjects segmented the text, indicated by pausing, did not always coincide with traditional grammatical boundaries, for example in the case of restrictive relative clauses, and subject and complement clauses. Schilperoord and Verhagen argued instead that the way in which the speaker makes conceptual links between clauses drives their segmentation of the text (see also Verhagen, 2001). For the restrictive relative clause, “if a constituent of a matrix- clause A is conceptually dependent on the contents of a subordinate clause B, then B is not a separate discourse segment” (Schilperoord & Verhagen 1998, p. 150) (for an example see 3 and 4 below). This echoes Langacker’s (1991) observation that for restrictive clauses the speaker can only conceptualise the referent in the matrix structure once they know the contents of the relative clause. This means that whether or not the matrix clause can be considered to be an independent unit of processing depends on the nature of the subordinate clause. To illustrate this, compare the texts in examples 3 and 4, below:

3. These schools all appear to have relatively many students who grew up in culturally deprived families
4. They shouted at the waiter, who so far did not seem to have noticed them.

For the restrictive clause in (3) we depend on the contents of the relative clause (“who...”) to understand the referent of ‘students’, because it provides necessary information. For the non-restrictive

---

1 Examples taken from Lieven et al.’s (2003) analysis of Annie’s data.
2 Examples 3 and 4 both taken from Verhagen, 2001
clause in (4) we can conceptualise the referent of ‘the waiter’ independently of the relative clause, because the relative clause simply provides extra information.

Schilperoord and Verhagen’s model provides a usage-based alternative to formal grammatical segmentation in speech because the linking of clauses is based on what they actually mean to the speaker rather than on abstract grammatical rules that are independent of lexicon and semantics. Moreover, the evidence suggests that the conceptual boundaries constitute a psychologically real unit of speech processing that traditional grammatical boundaries do not recognize. Segmenting a text according to conceptual conditions predicts the segmentation of text in spoken Dutch as measured by the good correlation between segments and pausing patterns (Schilperoord, 1996, 1997).

Further evidence that is consistent with Schilperoord and Verhagen’s theory is provided by Gee and Grosjean (1983). These authors did not look directly at functionally derived units; their units of analysis were ‘prosodic bundles’ that were derived from an algorithm based on formal linguistic principles, but importantly these units did not correlate with syntactic boundaries. In fact they closely approximated the ‘phonological word’ unit used by Howell and colleagues (discussed in section 5) that is made up of a content word surrounded by function words. Gee and Grosjean found that that pausing precedes a function-content word unit but hardly occurs at all at the boundary between the function and content words. The study can be interpreted, then, as showing that speakers pause before they embark on one of these function-content word units, and doing so avoids the risk of repetition of function words. Related to the evidence from Gee and Grosjean’s (1983) work, Pinker (1995) gave examples showing that pauses do not always occur at a major (formal) syntactic boundary. In the examples used to illustrate this, the pauses do, however, all occur at phonological word boundaries. So, again, this is consistent with speakers pausing prior to a planning unit that is not derived from formal syntax.

The naturalistic data are backed up by some experimental evidence. Vogel Sosa and MacFarlane (2002) used a word-monitoring paradigm to compare the reaction times of adult English speakers for recognizing the function word ‘of’ in frequent versus infrequent collocational contexts. They found that speakers responded slower to ‘of’ when it occurred in a highly frequent collocation, such as ‘kind of’ compared to a lower frequency collocation like ‘sense of’. This is consistent with the UB theory that words that very often occur together in connected speech are stored as a single unit and accessed holistically, as suggested by the UB model. To recognise ‘of’ in a frequent collocation would require decomposing the holistic unit into its individual parts, similar to the morphological decomposition required to access the past tense morpheme after hearing ‘walked’. The words in the holistic unit would only be recognized via their connections with other instances of the component parts ‘kind’ and ‘of’. The less frequent collocations would not be fused into units and so the recognition of ‘of’ would be direct, and hence faster. Though the data from Vogel Sosa and MacFarlane focus more on lexical access than syntactic processing, they do provide some preliminary evidence that UB units are relevant in real time speech processing. They suggest that speakers do not plan their speech in segments based on the units of formal linguistics, but rather in segments that derive from the input.

To summarize, UB theories predict that we can expect words to be stuttered at processing boundaries and these may not be the traditional grammatical boundaries of formal linguistics. Notably, the boundaries will vary over developmental time as children build up more abstract representations. The problem is that because the UB approach is still young, it has tended to focus mainly on the nature of the underlying grammatical representations and little is yet known about how the different routes to production (direct and via abstract units) interact online. Though the evidence discussed above is consistent with the proposal, there is little evidence that directly assesses what the alternative units of processing may be. The neglect is most noticeable in the child language literature, where no real theory has yet been provided from the UB perspective to explain the online mechanisms involved in children’s speech production. In other areas of study, however, the interest in early language production and fluency has been strong. There has been much debate about what factors affect the location of disfluencies in speech (e.g. grammatical complexity) and the specific nature of the mechanisms of language production that are entailed in this. We will turn now to a discussion of the existing literature surrounding these topics, before returning to the UB approach to explore how the two fields may be able to contribute to one another.

4. The Wider Relationship between Linguistic Representations and Fluency in Speech Production

The relationship between natural language factors and stuttering is widely researched in children who stutter (Wingate, 1988; 2002). Generally speaking (and on first sight somewhat counter-intuitively), fluency problems are evident on function words (Bloodstein & Gantwerk, 1967; Bloodstein & Grossman, 1981) and involve repetition of the whole of these words (Conture, 1990), e.g. ‘at at at school’. Young speakers seem not to be influenced by the phonetic structure of the words (e.g. whether...
the word contains a consonant string or not, or whether the consonants in a word have manners that are
difficult for a child to produce such as fricatives and laterals) (Howell, Au-Yeung & Sackin, 2000). In
contrast, fluency problems in adults who stutter are evident on content words (Howell, Au-Yeung &
Sackin, 1999) and the disfluency often involves the first part of these words (Conture, 1990), e.g. ‘at
ssssssschool’. Their fluency is affected when the content word has properties that make words difficult
to acquire for children, e.g. ‘school’ would be difficult because it has a consonant string (sk) containing
a late emerging consonant (s) in word-initial position (Howell, Au-Yeung & Sackin, 1999).

Different approaches have been taken to explaining why childhood stuttering is anomalous. Wingate
(2002) argued that the childhood pattern is just normal nonfluency, not stuttering (which would also
explain why there is a high rate of recovery from childhood ‘stuttering’). The majority of authorities
maintain that there is developmental change (Conture, 1990; Howell, 2004). Howell goes further and
argues that the different patterns of stuttering reflect contextual influences. Word repetitions are
stuttering-like disfluencies that precede difficult words. Howell argues that the childhood form of
disfluency is associated with getting the words ready in time (see later for further details).

The empirical evidence concerning the relationship between syntax and fluency remains equivocal.
The joint questions of whether fluent children (children who do not stutter) (CWNS) use more complex
syntactic structures than children who stutter (CWS) and whether CWS have deficient syntactic
capacity have yielded contradictory findings. Most authors agree that children are more likely to
produce disfluencies on utterances that are more syntactically complex (see Bloodstein, 1995; Karmiol,
1995; and Ratner, 1997; for reviews) but there is wide variation as to how this relationship should be
interpreted. Many authors hold that stuttering is the external symptom of the difficulty that CWS
experience with syntactic processing, and that this is evidenced by a correlation between syntactic
complexity and disfluency rate (e.g. Blood & Hood, 1978; Bernstein, 1981; Bruten & Hedge, 1984;
Gordon & Luper, 1989; Gaines, Runyan & Meyers, 1991; and see Karmiol, 1995). Others argue that the
interpretation of these conclusions is a less straightforward matter. Yaruss (1999) also found that
disfluency was related to syntactic complexity, but showed through logistic regression that this was a
poorer predictor of stuttering than was length of utterance, and that neither of these measures was a
particularly strong predictor. He argued that these factors cannot adequately account for stuttering by
themselves. Logan and Conture (1997) argue that the relationship they found between stuttering and a
higher number of clausal constituents in the utterances of CWS (regardless of length of utterance) does
not necessarily reflect syntactic processes, but could instead reflect prosodic planning. An argument has
been made that CWS are less syntactically able than CWNS. Some studies show that CWS use less
complex utterances than CWNS (e.g. Wall, 1980; Wall, Starkweather & Cairns, 1981; Howell & Au-
Yeung, 1995) but others have found no difference between CWS and CWNS on measures like the
Developmental Sentence Analysis (Westby, 1974) and some even suggest a relative level of syntactic
precocity in CWS when compared to norms (e.g. Ratner & Sih, 1987; Watkins & Yairi, 1997). The
relationship is indeed complex. Westby (1974) found that CWS make more grammatical errors than
CWNS but Yaruss (1999) found no difference in grammatical accuracy of the stuttered and fluent
utterances of CWS.

Some of the confusion regarding the relationship between syntax and fluency results from the wide
variation in how different studies appraise syntactic complexity. Several different measures have been
used in various combinations by investigators to classify children’s productions in naturalistic data.
They include the number of clausal constituents a sentence contains (e.g. Logan & Conture, 1997), the
number of clauses in an utterance (Yaruss, 1999; Wall, 1980), the amount of embedding in an utterance
(Kadi-Hanifi & Howell, 1992) and the earliest age at which different structures are produced by
children (e.g. passives and negatives occur later than actives and declaratives, so are considered harder)
(e.g. Silverman & Bernstein-Ratner, 1997; Yaruss, 1999). Methods of assessing children’s underlying
knowledge have also been used, including the Reception of Syntax Test (ROST) (Howell, Davis & Au-
Yeung, 2003) and elicited imitation paradigms wherein children try to reproduce sentences that
comprise various levels of syntactic complexity (Silverman & Bernstein-Ratner, 1997).

The methodological disparity inherent in the literature can partly explain the confusing pattern of
results but there is, in fact, a hidden assumption that is shared by all of the methods mentioned. They all
characterize syntax in the traditional sense, with all exemplars of a general syntactic category being
treated with equivalence (a possible exception being Kadi-Hanifi & Howell, 1992, who also drew on
semantic factors). None of the methods examined systematically the relationship between the syntactic
frame that is used by a child and the lexical content therein. The importance of this should be clear,
especially for young children (2-3 years), from section 2 where we considered the nature of children’s
early syntactic knowledge as revealed by the UB approach. For example, if passives are classified as
more complex than the actives because they emerge at a later stage of development in English, their
lexical composition still cannot be ignored. If a child produces a passive utterance, we have already seen
that we cannot be sure how they have gone about doing so. A passive could be constructed using a
single operation to modify a highly concrete schema like ‘The W got broken’ (where W means a slot that contains a variable word). This would actually be easier to construct than a more creatively produced active construction, as when a child inserts words and rearranges them to form ‘Mummy’s pushing me now’ from ‘Mummy’s trying to push me’. This may be the case even when utterance length is the same. As shown, a classification of children’s utterances based on adult-like abstract categories does not necessarily represent accurately the capacities of young children. Instead, a more appropriate and representational measure would be how creative the utterance is, for instance how abstract the component parts are (e.g. is the construction used with only one verb or many different ones?) and the number of manipulations needed to reach the present utterance from previous ones (i.e. compare one structure with all preceding instances).

A related and crucial issue is the inability of absolute measures to take account of the changing nature of children’s syntactic knowledge across age. For example a measure that classifies all passives as complex will attain different results for a child as they grow older, because early passives are likely to actually be more simply produced from lexical schemas, whereas later instances are likely to be produced from more abstract underlying constructions. The UB approach has assembled a large body of evidence about the way in which children’s syntactic knowledge changes as they grow older and this knowledge is ripe for application to fluency research. To exploit this potential, we need first to find a suitable model of the processes involved in speech production in real time. This will allow us to link together the UB approach insights into the units of linguistic representation with the processing mechanisms that act on these to transform them into the spoken speech signal.

5. The EXPLAN Model of Speech Production (Howell & Au-Yeung, 2002)

The EXPLAN speech production model of Howell and Au-Yeung (2002) is unique in being the only model to be explicitly developmental, and to explain the high incidence of the anomalous stuttering on function words in early development. EXPLAN arose out of the work of Howell, Au-Yeung and Sackin (1999) who investigated the relation between stuttering on function and content words within a contextual unit, the phonological word, which specifies the extent of units that incorporate the two types of word (see also Au-Yeung, Howell & Pilgrim, 1998, who first introduced phonological words into analysis of stuttered speech).

Phonological words (PWs) as defined by Au-Yeung et al. (1998) have an obligatory content word and a variable number, from zero up, of function words preceding and following it. An example is ‘I split it’ which has one function word before, and one after, the content word (the verb ‘split’). The function words in a PW are associated with their content word by sense unit rules (i.e. the function word has to be semantically related to its content word). Three properties of stuttering are seen in PWs that have function words before and after the content word. First, stuttering on function words, on the vast majority of occasions, occurs on those that precede the content word (on ‘I’ in the preceding example) (Au-Yeung et al., 1998). Second, stuttering occurs either on the function word or words that precede the content word or the content word itself, not both (Howell et al., 1999) – you see ‘I, I, I split it’ or ‘I sssplit it’ commonly but rarely see ‘I, I, I sssplit it’. Third, the tendency to stutter more on content words as age increases is associated with a corresponding decrease in stuttering on initial function words (Howell et al., 1999). A fourth important feature, not specifically about the distribution of stuttering, is that the type of stuttering on function words tends to involve hesitation around, or repetition of, the whole function word whereas stuttering on content words typically involves difficulties producing the first part of these words as in part-word repetitions (‘s... split’) or prolongations (‘sssplit’) (Conture 1990). Another relevant finding is the work by Gee and Grosjean (1983) mentioned above, showing that pausing precedes a function-content word unit, but hardly at all at the boundary between the function and content words.

The EXPLAN model offers an account of all of these findings on the assumption that at the root of all stuttered output is the difficult word (that for English is usually a content word). In the EXPLAN model, fluency control problems arise because the complex content words can take too long to generate for the context in which they need to be produced. In a connected stretch of speech, like the PW in the earlier example, the plan for the content word may not be ready for execution immediately after the initial function word has been completed and, therefore, the whole content word cannot be executed. The speaker can do one of two things to deal with this problem: First, the speaker can interrupt speech by pausing or re-executing the words that precede the word that is not ready (i.e. the initial function word or words). Pausing or repeating gains time to complete the plan of the content word, but they will only work when these disfluencies occur on the initial function word in an example like ‘I split it’ (accounting for the first feature noted above). To be concrete, ‘I split it it’ occurs rarely because repetition of ‘it’ cannot gain more time for planning a content word that precedes it (‘split’ in this case). When word repetition or hesitation around the initial function words occurs, it provides more time that, in turn, prevents stuttering on the content word. This explains why in any particular disfluent PW,
stuttering happens in an either-or fashion on initial function or content words (the second feature noted earlier). Finally, the repeated function words are produced in their entirety as their plan is complete.

The second possibility for the speaker is to start the utterance after the initial function word has been produced, even though its plan is not complete. The part of the plan that would be available is the initial part (assuming speech is generated left to right). There is then no function word repetition or hesitation, rather stuttering occurs on the content word. If the plan runs out (feature two), the first part is all that can be produced as that is all that is available (feature four). To account for the developmental changes, Howell and Au-Yeung (2002) assumed that as speakers get older, they change from responding to situations where the plan cannot be generated in time using function word repetition to producing parts of content words (feature three). The issue as to what underlies this change has not been worked on to date (though see Howell, 2004, for some hypotheses).

6. The Relationship between EXPLAN and the UB Approach

EXPLAN provides us with a developmentally orientated model of online speech production that takes account of the planning context of the stuttered units. However, the PW as the unit of speech production is not derived empirically but from phonological theory (Selkirk, 1984). Naturalistic evidence that adopts it as a unit of analysis fits well with the EXPLAN model but is this because the unit is psychologically real or does it simply correlate with other units that are, namely those identified by the UB approach? Some researchers would argue that the difference between the UB units and the PW resides in a distinction that can be drawn between ‘linguistic’ and ‘speech output’ representations. This is a matter of theoretical perspective but at least in the case of children there is no empirical evidence to show that this distinction is necessary or justified.

The PW is an abstract component of generative phonological theory in the same way that traditional clausal boundaries and constituents are abstract components of generative syntactic theory. In fact, the need for generative phonology only arose because generative syntax failed to explain why the phonological and prosodic groupings of spoken language do not always match up with syntactic boundaries. Formal linguistics developed phonological theory as a further layer in the language production process to account for these patterns (Jackendoff, 2002). In fact, viewed as a whole, the generative model is rather clumsy and, in its attempt to fit theory to data by addition rather than modification, it loses the elegance that its early supporters admired. UB models are more parsimonious and maintain that syntax is not completely isolated and is connected to semantic and phonological knowledge, even in adults. At least until the evidence is gathered to resolve the issue, it seems premature to assume that children’s phonological knowledge is any more abstract than their syntactic knowledge. The slot-and-frame schemas provide a useful alternative unit of analysis for children’s fluency because their early phonological knowledge may well be tied to these if it is not abstract. It is worth mentioning that the authors do not reject the possibility that there are influences other than at the syntactic level that could influence fluency (e.g. prosodic planning), but the most detailed UB work to date with children concerns syntactic patterns. The argument is that until more is known about other forms of processing, the slot-and-frame schemas constitute a more empirically justified starting point for researching the planning units relevant to fluency in childhood than do the units of the generative approach.

How might schemas and operations interact online to affect the fluency of speech production? Naturalistic analysis should be expected to reveal a correlation between patterns of fluency and the location of schema boundaries as Howell et al. (1999) found with PW (c.f. Au-Yeung, Gomez & Howell, 2003, for Spanish; Dworzynski, Howell & Natke, 2003, for German). To date, the distribution of disfluencies in schemas has not been investigated, either with CWNS or CWS. It would be informative to investigate the validity of this hypothesised correlation for several reasons: 1) It would test directly the strength of the schema theory of early grammar as a psychologically real mechanism of language production; 2) A schema is an empirically identifiable mode of linguistic organisation for children and therefore constitutes an interesting candidate for investigating how fluency relates to syntax in a psychologically real sense. As mentioned above, the key point is that more disfluency would be expected at sensitive planning moments. The next section will make concrete suggestions for future research along these lines.

7. Suggestions for Future Fluency Research within the UB Approach

Through the course of this article, we have explored how the insights from the UB approach to language acquisition could relate to the existing research on fluency behaviour across development. There is much scope for reciprocal benefit by combining the efforts of researchers in child language and fluency. The UB approach has a tradition of embracing research from diverse disciplines (e.g. linguistics, psychology, neural net modeling) and it is hoped by the authors of the current paper that this
process of integration will continue in future by encompassing the area of fluency research and online speech production mechanisms. Several of the possible directions are suggested below.

The first logical step towards tying the two fields together would be to assess whether the units of speech (schemas) that Lieven et al. (2003) identified in Annie’s speech constitute online planning units in speech production for Annie herself. Though Annie is a fluent child, CWS and CWNS show the same forms of disfluency (see section 4), so the units she used would be a useful basis for searching for similar planning units in the speech of CWS. This would entail investigating whether the boundaries of the schemas correlated with her patterns of fluency, namely whether moments of disfluency coincided with the point in time at which an operation would be required to fill a slot (i.e. immediately before the slot word). The initial signs are that Annie did show more disfluency at 3;0 than at 2;0 but there is a real need to study this systematically. If the operations on schemas were empirically verified as planning units in speech production, the analysis of Lieven and colleagues could be used to study CWS. Around 25-30 hours of data would be required from a CWS so that the utterances in the final session could be compared with the earlier utterances. The operations and schemas that were identified in this way would be expected to correlate with the fluency patterns of the CWS.

This type of research would consolidate and extend the UB findings by providing a different kind of evidence for the existence of schemas and by investigating their online role. It would also improve fluency research by providing an empirically justified unit of analysis rather than the existing ones that are based on traditional generative theory, which does not match the data on child language. To date, the PW adopted by Howell and colleagues has provided the best fit with the evidence on fluency compared with traditional grammatical units, but this is a unit motivated by phonological theory rather than being empirically derived. The ‘slot-and-frame’ operations of Lieven et al. (2003) are compatible with the existing data because most of the abstracted ‘slots’ in Lieven et al.’s schemas from Annie’s data are content words that are preceded by function words at the end of the ‘frame’, for example Annie’s ‘Where’s the X?’ could become ‘Where’s the bus!’ or ‘Where’s the cat?’ and her ‘I want a Y’ could become ‘I want a toastie’ or ‘I want a biscuit’. There is also work by Strenstrom and Svartvik (1994) on fluent adult English speakers that showed more fluency breakdown occurred on pronouns that were produced before verbs than after verbs (3.39% and 0.14% were repeated, respectively). The finding is consistent with the hypothesis that the PW is the unit of speech planning (i.e. predicts that function words will be repeated prior to content words rather than after them) but also that low-scope schemas are the units of speech planning, especially as these are often based around pronouns (see section 2). It would be interesting to compare traditional grammatical segmentation with PW segmentation and UB schema segmentation in terms of how well they predict the planning unit boundaries of children’s speech as determined by their patterns of fluency and disfluency.

Another interesting issue that has been mentioned briefly is whether there are any differences in the syntactic knowledge or abilities of CWS compared with CWNS. The tools of the UB approach could be used to address this by assessing the syntactic knowledge of CWS (e.g. the novel word paradigm) and the nature of the naturalistic productions of CWS (e.g. how much abstraction versus lexical specificity is evident at 2 years compared with CWNS). The need to study this is supported by Silverman and Bernstein-Ratner’s (2002) recent finding that there may be differences between CWS and CWNS in terms of the variety of lexical items they use in their speech, with CWS exhibiting less lexical variety than CWNS. It now remains to link this rather gross measure to the specific structures the children use, to see whether CWS may lag behind in the process of forming abstractions and so endure a more prolonged ‘trade-off’ period during which speech disfluencies occur. One problem with providing a rich data set from a young child who stutters is that data tend to be available at a later age, as it is only then that the child is identified as being a child who stutters. Given that UB theories of grammar propose that schemas or constructions underlie grammar right up to adulthood, however, the answer would be to adapt the analysis to search for more abstract level schemas than Lieven et al. (2003) did. Once such analyses were available, they would present a rich seam of data, but in the absence of a rich enough data set, an alternative less labour-intensive solution is possible. Carefully selected ‘probe’ items could be used. Instead of performing a full distributional analysis to determine the candidates for planning units, the existing evidence from adult data could be used to provide high frequency probe items such as Bybee and Scheibman’s (1999) ‘don’t’ (mentioned earlier) to assess whether these correlated with patterns of fluency. Again, this type of research would both help to test the UB approach and shed further light on stuttering.

One more area in which the UB approach would benefit fluency research is in assessing the underlying phonological representations of CWS (though this has not been the focus of the current article). This would be a method of exploring the discrepancy between the fluency research which argues that the relevant planning unit for fluency behaviour is the PW, and the UB account, which suggests that low-scope schemas may be the units of planning in speech production for children. The issue could be approached empirically by examining lexical specificity in relation to phonology in a
similar way to that which UB researchers have done for syntax. If the lexical content of certain prosodic patterns in children’s speech varies, then it could be claimed that the child has abstract prosodic knowledge, whereas if the same phrases or schemas always occurred with the same prosody, then the child could be argued to have lexically-specific prosodic knowledge instead. The UB approach would predict that early in life these prosodic and syntactic features are stored together in a schema, and only later do they become useable separately. A different issue concerning phonology would be to establish whether CWS have accurate and intact phonological representations compared to CWNS. A priming paradigm could be used, for example, to test whether CWS and CWNS who hear ‘ba’ would be primed by this to name a picture of a banana more quickly. If so, then we could conclude that both CWS and CWNS possessed the full representation of banana and CWS do not stutter as a consequence of poor phonological representations. On top of the priming effect, if CWS were found to name the words slower than CWNS, then it could be argued that they encode phonological representations more slowly (as suggested by Kolk and Postma’s 1997 ‘Covert Repair Hypothesis’ account of stuttering), though this would not be the prediction of the EXPLAN model discussed earlier. Thus there are many ways in which the tools of the UB approach could be of benefit to fluency researchers interested in phonology.

To conclude, the UB approach to child language development and research on fluency have much to contribute to one another. The ideas presented here are by no means an exhaustive representation of what is possible. It would be very productive for both fields if researchers were to collaborate and combine the insights of both. At a practical level, deeper understanding of fluency development in general is essential for the proper management of the disorder in childhood and throughout life.

Acknowledgement. The first author was supported by the Wellcome Trust. The second author is grateful to Heike Behrens, Stephanie Brosda and Michael Tomasello for initial discussions on this topic.

References


