Bilingualism: Language and Cognition 1 (1998) 1–20 © 1998 Cambridge University Press

Developmental dynamics in L1 and L2 acquisition: Processability Theory and generative entrenchment*

MANFRED PIENEMANN
Australian National University

This paper has two major objectives: (1) to summarise Processability Theory, a processing-oriented approach to explaining language development and (2) to utilise this theory in the comparison of development in L1 and L2 acquisition. Proponents of the Fundamental Difference Hypothesis (between L1 and L2) assume that L1 development can be explained with reference to Universal Grammar (UG) which, in their view, is inaccessible to L2 learners. Instead, they claim that a second language develops on the basis of language processing strategies.

I will show that the fundamentally different developmental paths inherent in first and second language acquisition can both be explained on the basis of the same language processing mechanics (as specified in Processability Theory). I will demonstrate that the developmental differences between L1 and L2 are caused by the qualitatively different early structural hypotheses which propagate through the acquisition process. The concept of "propagation of structural features" will be viewed as "generative entrenchment," a logical-mathematical concept, which has proved to be highly productive in examining other kinds of developmental processes.

A brief sketch of Processability Theory

The wider context

Learnability is defined as a purely logico-mathematical problem (e.g. Berwick & Weinberg, 1984). Such a perspective ignores the fact that this problem has to be solved, not by an unconstrained computational device, but by a mind that operates within human psychological constraints.

In this section I will sketch out a theory which adds to learnability theory the perspective of processability, that is, "Processability Theory" (Pienemann, in press). In my view, the logico-mathematical hypothesis space in which the learner operates is further constrained by the architecture of human language processing. Structural options that may be formally possible will be produced by the language learner only if those processing resources are available that are needed to carry out, within the given minimal time frame, those computations required for the processing of the structure in question. Once we can spell out the sequence in which language processing routines develop in the learner, we can delineate those grammars that are processable at different points of development.

The architecture of human language processing therefore forms the basis for Processability Theory. In this perspective the language processor is seen with Kaplan and Bresnan (1982) as the computational routines that operate on (but are separate from) the native speaker's linguistic knowledge. Processability Theory primarily deals with the nature of those computational routines and the sequence in which they become available to the learner. It will be argued that language acquisition incorporates as one essential component the gradual acquisition of those very computational routines. In other words, the task of acquiring a language includes the acquisition of the *procedural skills* needed for the processing of the language. It follows from this that the sequence in which the target language (TL)[1] unfolds in the learner is determined by the sequence in which processing routines develop that are needed to handle the TL's components.

In the rationalist tradition, learnability analyses have in the past been based on four components that

* I would like to thank Avery Andrews, Bruno Di Biase, Gisela Håkansson, Malcolm Johnston, Jürgen Meisel, Bonnie Schwartz and four anonymous reviewers for their invaluable input into this paper. Any remaining errors are mine. I gratefully acknowledge the financial assistance from the Australian Research Council which I received for the research presented in this paper.

[1] The term "target language" is used here to cover L1 and L2.

Address for correspondence
Linguistics Program, Department of Modern European Languages, Australian National University, ACT 0200, Australia
E-mail: manfred.pienemann@anu.edu.au

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

must be specified in any learnability theory (e.g. Wexler and Culicover, 1980; Pinker, 1979):

(1) the target grammar,
(2) the data input to the learner,
(3) the learning device that must acquire that grammar, and
(4) the initial state.

The idea behind this is that a learnability theory must specify how a learner develops from an initial state to the target grammar by means of the available input and the given learning device.[2]

The rationale for assuming these components is rooted in the way in which learnability theory has been formulated in response to the "logical problem" in language acquisition (cf. Wexler, 1982). The logical problem basically describes the following paradox: children acquire in a relatively short period of time and on the basis of limited linguistic input the basic principles of their native language, although it is assumed that many of these principles cannot be inferred from the observations made by the learner.

It has been noted by several rationalist researchers (e.g. Felix, 1984, 1991; Clahsen, 1992; Gregg, 1996) that besides linguistic knowledge, a theory of language acquisition must also explain what causes the development of the TL to follow a describable route. This explanatory issue has been referred to it as the "developmental problem" (Felix, 1984).

My fundamental point is that recourse needs to be made to key psychological aspects of human language processing in order to account for the developmental problem, because describable developmental routes result, at least in part, from the architecture of the human language processor. For linguistic hypotheses to be transformed into executable procedural knowledge (i.e. a certain processing skill), the processor needs to have the capacity of processing those hypotheses.

In other words, Processability Theory focuses solely on the developmental problem as an explanatory issue; it is not designed to contribute anything to the question of the innate or learnt *origin* of linguistic knowledge or the inferential processes by which linguistic input is converted into linguistic knowledge. Instead, it is the objective of Processability Theory to determine the sequence in which procedural skills develop in the learner.

[2] It has been noted that the potential components of a theory of learnability interact: the stronger the first, the weaker the second and vice versa or, in the words of Bates, MacWhinney, and Smith (1982):

> . . . As Pinker (1979) and Braine (1978) both noted, the Wexler and Culicover conclusion is not the only one that can be reached with learnability analysis. The strength of their (W & C's) fourth parameter (i.e. innate hypotheses) is required only because of the value they have assigned to the other three. (p. 15)

Key psychological factors in language processing

It is the aim of this paper to hypothesise, on the basis of the general architecture of the language processor, a universal hierarchy of processing resources which can be related to the requirements of the specific procedural skills needed to process the TL. In this way, predictions can be made for language development which can be tested empirically. Before I describe the architecture of the hypothesised hierarchy of processability and the model of language production on which it is based, it will be useful to sketch out briefly a number of key psychological factors in language processing in order to characterise the processing environment within which the learning of language takes place. The framework that follows from this will then be used to establish a hierarchy of processing prerequisites. The formal system adopted for the description of the processing hierarchy will be shown to reflect key properties of language processing.

The view on language production followed in this paper is largely that described by Levelt (1989), which overlaps to some extent with the computational model of Kempen and Hoenkamp (1987) which emulates much of Merrill Garrett's work (e.g. Garrett, 1976, 1980, 1982) on which the corresponding section of Levelt's model is based. The basic premises of that view are the following:

(1) processing components are relatively autonomous specialists which operate largely automatically;
(2) processing is incremental;
(3) the output of the processor is linear, while it may not be mapped onto the underlying meaning in a linear way;
(4) grammatical processing has access to a grammatical memory store.

Below I will briefly describe each of these premises and highlight some of the empirical research supporting them.

Processing components are relatively autonomous specialists which operate largely automatically

The appeal of this proposition is that it can account for the *speed* with which language is processed. Levelt shows that adopting the assumption that processing is neither autonomous nor automatic leads to serious problems. If the processing compo-

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

nents were not autonomous, exchange of information between all processing components (or levels of processing) would have to be co-ordinated by a central control and processing would have to be executed in a serial manner which would lead to extremely slow lock-step style processing. However, not only has it been shown that language processing is much faster than would be predicted by such a model, but empirical evidence also shows that different processing components exchange information in a parallel (i.e. non-serial) manner (Levelt, 1989; Engelkamp and Zimmer, 1983; Sridhar, 1988).

In addition, a central control would imply that the operation of the processing components is consciously attended to, while empirical studies have shown that grammatical information is normally not attended to and can only be memorised if attention is focused on it (Bock, 1978; Kintsch, 1974; Garman, 1990).

Autonomous specialist processing components can further be characterised as processing devices which are able to accept and pass on only information of a highly specific nature, for instance, only information which is concerning NPs. The advantage of such task-specificity is a gain in processing speed, since non-attended processes can be executed in parallel. Thus the notion of task-specificity is in line with the observation that grammatical information can only be memorised if it is attended to. Below it will become clear that the notion of automatic unattended processing is closely related to the nature of processing resources.

Processing is incremental

This premise basically says that "... the next processor can start working on the still-incomplete output of the current processor ..." (Levelt, 1989, 24). The idea is that the surface lexico-grammatical form is gradually being constructed while conceptualisation is still ongoing. This feature was highlighted for the comprehension system in a study by Marslen-Wilson and Tyler (1980) which demonstrated that in on-line processing semantic representations are being constructed by the comprehension system before grammatical structures have been "worked off."

Incremental processing is a feature of human language processing. One important implication of this feature for the structure of processing algorithms is that in order to be in line with human processing they must be able to cope with non-linear linguistic form without much "look ahead" (Levelt, 1989). In other words every processing component can "see" only a small section of the current processing event rather than having the complete event displayed.

Incremental processing therefore necessitates the use of storage facilities to allow for non-linearity in the matching of underlying meaning onto surface form. This has important implications for the concepts which will be developed below in relation to the acquisition of language. We will therefore briefly look at the interrelation between non-linearity and memory.

The output of the processor is linear, while it may not be mapped onto the underlying meaning in a linear way

One case of non-linearity is the relationship between the natural sequence of events and the order of clauses. As Levelt (1983) points out, propositions do not necessarily have to be produced in language in the natural order of events. Consider the following example:

(1) Before the man rode off, he mounted his horse.

In this example the event described in the second clause happens before the one described in the first clause. In order to produce such sentences then, the speaker has to store one proposition in memory.

There are similar *linearisation problems* (Levelt, 1983) which operate at the morpho-syntactic level. Such cases involves the storage of grammatical information. One such example is subject-verb agreement. Consider the following example:

(2) She gives him a book.

The insertion of the verbal agreement marker crucially hinges on the storage of grammatical information which is created before the verb is produced, namely person and number marking of the subject. Note that the nature of the information held in memory is not the same in this example as it was in the above example. In the previous example it was propositional content which had to be stored, while in this example storage for grammatical information is needed.

Grammatical processing has access to a grammatical memory store

Both types of information need to be held only temporarily until incorporated into the generation of the message. However, the difference between the two types of information is this: grammatical information is highly specific, and (conscious or unconscious) attention to it is not necessary; one does not need to be aware of or control the fact that the information concerning "person" and "number" matches between the lexical entries of the verb and the grammatical subject. In fact, it is possible to attend to only a small number of such processes.

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

With the normal speed of language generation Working Memory would otherwise get "clogged up". On the other hand, attention must be focused on the propositional content, because it reflects the intended conceptualisation the speaker wants to express.

Working memory is the resource for temporary attentive processes which include conceptualising and monitoring (Baddeley, 1990; Broadbent, 1975). It has a highly limited capacity and is therefore not suitable to process great amounts of grammatical information at high speed. Levelt (1989) and other authors (e.g. Engelkamp, 1974) assume that grammatical information is held temporarily in the grammatical memory store which is highly task-specific and in which specialised grammatical processors can deposit information of a specific nature. In Kempen and Hoenkamp's (1987) Incremental Procedural Grammar, the locus of the grammatical buffer is the specialised procedures which process NPs, VPs, and so on.

One can see that the grammatical memory store is a requirement that arises from the automatic (i.e. inattentive) nature of grammatical processing: grammatical processors handle highly specific information which the grammatical memory store can hold temporarily. Empirical evidence for the different nature of the processing of propositional content and grammatical information comes, among other things, from the study of aphasia. Cooper and Zurif (1983), for instance, showed that in Broca's and, to a lesser extent, in Wernicke's aphasia lexical retrieval and semantic representation are functional while grammatical information cannot be processed. This is true for production as well as for comprehension.

As with other motor and cognitive skills, automatic processes in language production utilise what is known as "procedural knowledge" or "procedural memory", which is contrasted with "declarative knowledge/memory". The latter "... concerns everything that can be represented at a conscious level, and which groups together what Tulving (1990) called 'episodic' memory, and Penfield termed 'experiential' memory (Penfield and Roberts, 1959)" (Paradis, 1994, 395). There is ample empirical evidence from studies of amnesia and aphasia for the dissociation of procedural and declarative memory (Paradis, 1994), based on the patients' loss of ability to perform or to learn to perform certain tasks which can be defined according to the procedural-declarative distinction.

Incremental language generation

The process of incremental language generation as envisaged by Levelt (1989) and Kempen and Hoenkamp (1987) is exemplified in Figure 1, which illustrates some of the key processes involved in the generation of the example sentence "A child gives a cat to the mother". First of all, the concepts underlying this sentence are produced in the Conceptualiser. For the purpose of this paper I will ignore the internal structure of this component of language generation, except for several features of the output produced by the Conceptualiser.

In the example chosen in Figure 1, the conceptual material produced first activates the lemma CHILD in the lexicon. The lemma contains the category information N which calls the categorial procedure NP. This procedure can build the phrasal category in which N is head, i.e. NP. The categorial procedure inspects the conceptual material of the current iteration (the material currently being processed) for possible complements and specifiers and provides values for diacritic features, including those from the head of phrase. I will assume that the first referent is marked "- accessible". This ensures that the branch Det is attached to NP, the lemma for "A" is activated, and that the lexeme "a" is inserted. Functorisation Rules instigate the activation of free grammatical morphemes and the insertion of bound grammatical morphemes.

The above attachment of Det to the NP-node illustrates a key feature of the language production process, which is crucial in the context of language acquisition. The selection of the lemma for "A" partly depends on the value of a diacritic feature ("singular") of the head being checked against that of the targeted lemma. The value of the diacritic feature is 'stored' by the categorial procedure until it is checked against that of the modifier.

Our production process has proceeded to the point where the structure of a phrase has been created and the associated lemmata are activated. What is missing to make this the beginning of a continuous and fluent utterance is establishing a relation between the phrase and the rest of the intended message. This is accomplished by assigning a grammatical function to the newly created phrase. In fact, it is the categorial procedure itself that chooses its functional destination. This highlights the active nature of syntactic procedures.

Possible functional destinations are defined in a set of so-called Appointment Rules which are also language-specific. The default for NP procedures is "subject of S." However, this does not quite solve the problem of allowing the tree created so far to grow into a sentence and to make the production of the sentence continuous. What is missing is the attachment of the NP to a higher node. In the above example NP_{subj} calls the procedure S which accepts the calling NP as its subject and stores the diacritic

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

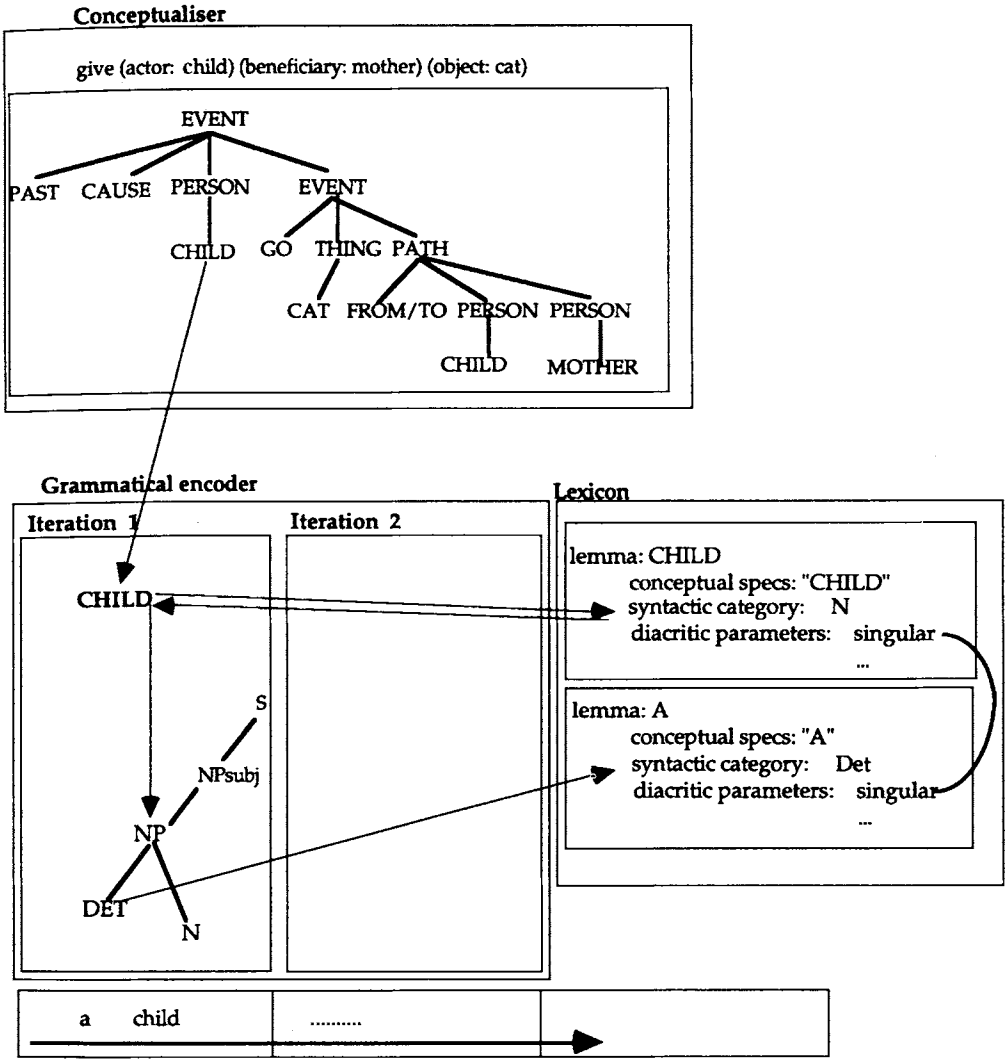


Figure 1. Incremental language generation.

features deposited in the NP, namely the values for "person" and "number".

The outcome of all of this is depicted by a tree structure in Figure 1. While this structure is produced and the associated lemmata are activated, the next conceptual fragment would have been processed in parallel and the output of the Formulator would have been delivered to the Articulator. This means that new conceptualisation occurs while the conceptual structure of the previous iteration is being produced. The whole process then moves on from iteration to iteration. This is what Kempen and Hoenkamp (1987) and Levelt (1989) mean by incremental production.

In the above summary of the process of grammatical encoding one aspect was left aside, namely word order. The definition of the acceptable set of word order constellations for configurational languages is carried out by Word Order Rules, which co-ordinate the assembly of phrasal subprocedures. I assume that for non-configurational languages grammatical roles can be specified directly from the semantic roles specified in the conceptual structure. And different procedures are assumed in the above model for the processing of matrix and subordinate clauses.

To recapitulate, in the incremental process of language generation, the following processing prerequisites are activated – among other things – in the following sequence:

(1) the lemma,
(2) the category procedure (lexical category of the lemma),
(3) the phrasal procedure (instigated by the category of the head),
(4) the S-procedure and the target language word order rules,
(5) the subordinate clause procedure – if applicable.

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

A hierarchy of processing resources

It is important to note that this account of grammatical processing and memory applies only to mature users of language, not to language learners. While even beginning second language learners can make recourse to the same *general* cognitive resources as mature native language users, they have to create language-specific processing routines. For L1 learners there are obviously no pre-existing procedures which can be transferred. L1 learners therefore have to develop all specific L1 procedures. Below I will generate hypotheses as to how language-specific processing routines develop, given the general architecture of the language processor.

In this context it is important to ensure that Levelt's model can, in principle, account for language processing in bilinguals, since second language acquisition will lead to a bilingual language processor. De Bot (1992) adapted Levelt's model to language production in bilinguals. Based on work by Paradis (1987) he shows that information about the specific language to be used is present in each part of the preverbal message, and this subsequently informs the selection of language-specific lexical items and of the language-specific routines in the Formulator. Drawing from Paradis's (1987) research, De Bot concludes that "... the speaker who speaks two closely related languages will for most part use the same procedural and lexical knowledge when speaking either of the two languages, while in the case of languages which are not related an appeal is made to much more language-specific knowledge." (De Bot, 1992, 9). De Bot further shows that Paradis's (1987) "Subset hypothesis" about the bilingual lexicon is in line with the overall design of Levelt's model. According to the subset hypothesis, the bilingual lexicon is a single storage system in which links between elements are enforced through continued use. This has the effect that links will be stronger between elements from one language. However, in the case of bilingual communities with a tendency for code-switching, links between elements from different languages may be similar to those in a monolingual lexicon.

De Bot (1992) demonstrates that the extended version of Levelt's model accounts for the majority of the additional requirements which have to be met by a language production model in a bilingual context. These include the following requirements. The two language systems concerned may be used quite separately from each other or in varying degrees of mixes (code-switching). The two systems may influence each other. Neither system will necessarily slow down in speech rate in comparison with a monolingual speaker, and the bilingual speaker may master the two (or more) systems to differing degrees.

Given the focus of Processability Theory on the Formulator, the key assumption from De Bot's work for the present context is that in all cases where the L2 is not closely related to the L1, different (language-specific) procedures have to be assumed. Based on our previous discussion, the following language-specific processing devices are the least L2 learners have to construct to acquire the L2 grammar:

- word order rules,
- syntactic procedures and their specific stores,
- diacritic features in the lexicon,
- the lexical category of lemmata,
- functorisation rules.

Obviously, word order rules are language-specific and there is no *a priori* way of knowing for the language learner how closely related L1 and L2 are. Learners therefore have to be equipped to bridge maximal typological gaps in their L2 acquisition. Diacritic features of lemmata contain items such as "tense", "number", "gender", "case" etc. Again it is obvious that the list of diacritic features varies from language to language.

Similarly, syntactic procedures that build constituent structures and store temporarily specific grammatical information such as diacritic features are not the same across languages. Given that diacritic features are language-specific and that these are stored in syntactic procedures, L1 procedures are not equipped to handle the specific storage task required by the L2.

The lexical category of lemmata may also vary from language to language. Again, the language learner is only fit to acquire any of the world's languages if he or she tests the lexical category for every new lexical item.

The reader will recall that Functorisation Rules instigate the activation of free and bound grammatical morphemes. And the same is true for grammatical morphemes as what was said about word order rules: these are language-specific and therefore have to be acquired with the L2.

Exchange of grammatical information

In other words, the L2 learner is initially unable to deposit information into syntactic procedures, because (1) the lexicon is not fully annotated, and, more importantly, (2) because even if the L1 annotation was transferred, the syntactic procedures have not specialised to hold the specific L2 syntactic information. For this reason one can predict that the

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

beginning learner is unable to produce any structures which rely on the exchange of specific L2 grammatical information using syntactic procedures, or in LFG terms the "unification" of lexical features.

One can expand on the principle of grammatical information exchange in line with the architecture of the Formulator. In Figure 1 above I illustrated the flow of grammatical information in the production process and the different temporary stores used in that process. One type of process is the use of grammatical information that proceeds without reliance on temporary storage. An example is the morphological marking of reference to time. The information about tense is contained in the verb lemma with the value "past" for the diacritic feature "tense'. This means that the diacritic feature in question is available in the same location where the morpheme for the marking of past (i.e. "-ed") has to occur and no information has to be deposited into any syntactic procedure to achieve this process. I call the resulting class of morphemes "lexical." Since lexical morphemes can be produced without phrasal procedures they can develop before phrasal procedures.

A second type of process is one in which the grammatical information is stored in a phrasal procedure. An example was given in Figure 1, namely the NP "a child". The lemma CHILD is marked "singular", and the value of this diacritic feature has to match that of the determiner. To achieve this, the lemma information for CHILD has to be deposited in the NP-procedure and kept there for the activation of the lemma "A".[3] In other words, this second type of morpheme is linguistically characterised by agreement between the head of phrase and another phrasal constituent. Its processing characteristic is that of the storage of diacritic features in phrasal procedures. We can infer that this type of process will become available to the language learner once phrasal procedures have been developed for the L2. I call this class of morphemes "phrasal".

There is one further type of process which involves the exchange of grammatical information, namely agreement between heads of different phrases as in subject-verb agreement. For SV agreement to occur, the diacritic features "third person" and "singular" have to be deposited in the S-procedure until utilised for the activation for the verb lemma. I call this class of morphemes "inter-phrasal". This process obviously requires the S-procedure to be fully developed, and it involves the exchange of grammatical information between phrases.

One can now see that the three classes of morphemes are based on different processes. Lexical morphemes do not require phrasal procedures while phrasal morphemes do. On the other hand, phrasal agreement will mostly occur in one and the same iteration. This is unlikely to apply to inter-phrasal agreement because of the incremental nature of language production. In other words, while the one phrase is being produced, the head of the agreeing phrase has not been conceptualised. This means that the relevant diacritic information cannot be stored in the phrasal procedure. Instead, it has to be stored in the S-procedure. However, in order to arrive there the functional destination of the phrase from which it originates has to be determined. This is carried out by a set of language-specific Appointment Rules as discussed above.

Principles of processability

The above brief exposition of the processability of different classes of morphemes may serve as an illustration of the fundamental principle of processability. Language-specific processing resources have to be acquired to make the processing of the TL possible. These processing resources are interrelated in two ways. (1) They feed into each other in the temporal event of language generation; i.e. one is utilised before the other. (2) The information processed and generated in one is required in the other. In this way these resources from a hierarchy. If one building block of the hierarchy is missing, the top cannot reached. The following processing resources were discussed above as part of the incremental language generation process:[4]

[3] Kempen and Hoenkamp conceptualise the matching of diacritic features as a process of "feature-copying" where features have a source and a destination. Vigliocco, Butterworth and Garrett (1996) demonstrate in cross-linguistic experiments on subject-verb agreement in Spanish and English that it is psychologically more plausible to view the matching of diacritic features as a process of "feature-merging" (or unification) in which both constituents involved in the matching process can derive features independently from conceptual structure. They show that the advantage of the merging process is that it can account for null-subject languages such as Spanish and for certain distributivity effects of conceptual structure on agreement phenomena. Vigliocco, Butterworth, and Garrett (1996) therefore assume a version of Incremental Procedural Grammar in which the feature copying mechanisms are replaced by feature merging mechanisms. In Lexical Functional Grammar this feature merging process is managed through the "unification" of features found in lexical entries. In Pienemann (in press) I adopt Vigliocco, Butterworth and Garrett's modification because of its psychological plausibility and because this creates a tighter fit with LFG.

[4] "Appointment rules" which were mentioned in the text above are not productive for the phenomena discussed in this paper. They do therefore not appear in this hierarchy.

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

- word/ lemma,
- category procedure (lexical category),
- Phrasal procedures (head),
- S-procedure and Word Order Rules,
- matrix/ subordinate clause.

These processing resources form a hierarchy which arises from the fact that the resource of each lower level is a prerequisite for the functioning of the higher level. A word needs to be added to the L2 lexicon before its grammatical category can be assigned. The grammatical category of a lemma is needed before a category procedure can be called. Only if the grammatical category of the head of phrase is assigned can the phrasal procedure be called. Only if a phrasal procedure has been completed and its value is returned can Appointment Rules determine the function of the phrase. And only if the function of the phrase has been determined can it be attached to the S-node and sentential information be stored in the S-procedure. And only if the latter has occurred can the target language word order be arranged. In other words, it is hypothesised that processing devices will be acquired in their sequence of activation in the production process. The reason for this assumption is that they constitute an implicational hierarchy in the production process. The implicational nature of the hierarchy would make it impossible for processing devices to develop before all other requisite devices have developed.

What happens when an element is missing in this implicational hierarchy? My hypothesis is that for the learner grammar the hierarchy will be cut off at the point of the missing processing device and the rest of the hierarchy will be replaced by a direct mapping of conceptual structures onto surface form as long as there are lemmata that match the conceptually instigated searches of the lexicon.[5] A point in case is the development of target language word order. Below I will demonstrate that the production of subject-verb inversion in German requires certain grammatical information to be exchanged between the grammatical subject and the verb. I will show that for such an exchange of grammatical information to be possible it has to be held in the S-procedure. I will show that as long as this information exchange is not enabled by the available processing prerequisites the learner is limited to the use of canonical sentence schemata. The only alterations that are possible at this point in development are ones that rely on additional general cognitive mechanisms such as perceptual saliency.

Indeed, the above principle for establishing an accessibility hierarchy of processing prerequisites was that of grammatical information exchange. This can be complemented by a non-linguistic principle, namely *perceptual salience.* Murdock (1962) established this principle through a number of studies which found persistent primacy and recency effects on the memorisation of any sequence of stimuli (cf. also Sridhar, 1988; Kintsch, 1974). The first and the last stimulus is more marked than the other stimuli and is persistently reproduced and learned better (cf. Kintsch, 1970). I will show below that saliency is a principle that is particularly productive in the learner's approximation to TL word order rules. With this addition the hierarchy of processing resources can be summarised as follows:

- word/lemma,
- category procedure (lexical category),
- Phrasal procedures (head),
- S-procedure and word order rules + saliency,
- S-procedure and word order rules – saliency,
- matrix/subordinate clause.

LFG and processability

If the above hierarchy is to be universally applicable to language acquisition, then it needs to be interpretable in relation to grammatical structures in individual languages. This is achieved by interpreting the processability hierarchy through a theory of grammar which is typologically and psychologically plausible. The theory of grammar I chose for this purpose is Lexical Functional Grammar (LFG), which shares three key features with Kempen and Hoenkamp's procedural account of language generation, namely (1) the assumption that grammars are lexically driven, (2) the assumption that functional annotations of phrases (e.g. "subject of") assume the status of primitives and (3) the mechanism of feature matching.

The key aspect of operationalising the above hierarchy of processing resources with LFG is the fact that the aspect of procedural memory inherent in Incremental Procedural Grammar (IPG) can be captured by the process of feature unification in LFG. Above I illustrated the aspect of procedural memory with SV agreement marking and I noted that for SV agreement to occur the diacritic features "third person" and "singular" have to be deposited in the S-procedure until it is matched with that of the verb entry. In LFG, this process is modelled by feature unification.[6] I will therefore use feature unification as

[5] Note that the sequence of generation of the preverbal message in the conceptualiser does not have to coincide with the constraints on word order.

[6] A proviso on this is that the procedures that underlie LFG cannot be understood to represent psychological procedures themselves. Instead, they can be considered a short-hand notation which

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

a comparative metric to evaluate the developmental level of inter-language (IL) forms.

Like Pinker (1984) and Levelt (1989), I use LFG as a convenient reference point which has been formalised and tested sufficiently to be practical for this purpose. The architecture of LFG coincides with most of the key points made above in relation to language processing. This does not mean, however, that I want to argue the case for LFG as the optimal theory in relation to language processing.

Second language development

The explanatory potential of Processability Theory will become clearer when the theory is applied to a particular case of language development. In this section I will apply Processability Theory to explaining the sequence found in the second language acquisition of German word order. This explanation will also constitute a crucial point of reference in the comparison of L1 and L2 development in the third and fourth sections below. The German L2 sequence is well-documented in the literature and has been the object of much of the debate on explaining developmental phenomena in Second Language Acquisition (SLA). The acquisitional sequence in question is based on a series of longitudinal and cross-sectional studies by the ZISA research group[7] (Clahsen, 1980; Clahsen, Meisel and Pienemann, 1983; Meisel, Clahsen and Pienemann, 1981; Pienemann, 1980, 1981). Similar findings have emerged in studies of the acquisition of German in formal contexts (Jansen, 1991; Pienemann, 1988; Westmoreland, 1983). In all cases, the basic sequence of acquisition can be summarised as follows:

Stage x = Canonical Order
die kinder spielen mim ball
"the children play with the ball"
Stage x + 1 = Adverb Preposing (ADV)
da kinder spielen
"there children play"
Stage x + 2 = Verb Separation (PART)
alle kinder muß die pause machen
"all children must the break have"
Stage x + 3 = INVERSION (INV)
dann hat sie wieder die knoch gebringt
"then has she again the bone bringed"
Stage x + 4 = Verb Final (V-END)
er sagt, daß er nach hause kommt
"he says that he home comes"

It should be noted that, in the process of L2 acquisition, the learner accumulates these rules. This means that the structure of a given IL can be described as the sum of all the rules the learner has acquired up to a certain point.

The application of Processability Theory to this sequence will occur in two steps: first a brief account of the word order phenomena concerned will be given in a somewhat simplified LFG framework, and then I will demonstrate that the process of feature unification required for each of the structures is predicted by the above hierarchy of processability.

In LFG possible word orders of languages are defined through c-structure rules (Kaplan and Bresnan, 1982, 175ff.).

(R1) $S \rightarrow NP_{subj}\ V\ (NP_{obj1})\ (NP_{obj2})$

(R1) allows for a basic SVO order (i.e. NP_{subj} V NP_{obj1} NP_{obj2}), as it occurs at stage x. The occurrence of wh-words, PPs and NPs in focus position, the characteristic of stage x+1, can be accounted for by (R2) which is adapted from Pinker (1984, 278):

$$(R2)\quad \bar{S} \rightarrow (XP) \qquad S$$
$$\left\{ \begin{array}{l} wh\ =_c + \\ adv\ =_c + \\ NP\ =_c + \\ PP\ =_c + \end{array} \right\}$$

The control equations in (R2) (e.g. wh $=_c+$) ensure that only wh-words, adverbs, NPs and PPs can occur in focus position. Standard German also allows lexical verbs to occur in this position, but such structures appear later than at x+1.[8] Also, note that at stage x+1 structures like "hat-ge-sag-en" (have-past-sag-infinitive) may occur. I analyse this as "PAST-PAST-V" and thus as a single verb entry, which can be inserted into the V-slot in (R1).

Hence (R2) allows for the possibility that the topic position becomes available separately to wh-words, adverbs, etc., because each of these categories requires separate control equations, which may be acquired individually.

The German "split verb" position (i.e. stage x+2

contains the necessary elements to relate structures to a hierarchy of processability. The LFG formalism is designed to be highly noncommittal as to when unifications are performed. They can be done incrementally, as each phrase is built, or at the end, when an entire c-structure has been constructed (see Maxwell and Kaplan (1995) for some discussion). Since Processability Theory assumes strict limits on grammatical memory, it would follow that unifications ought to be done as soon as possible.

[7] ZISA stands for "Zweitspracherwerb Italienischer und Spanischer Arbeiter" (Second Language Acquisition of Italian and Spanish Workers). The group was founded and directed by Jürgen Meisel at the University of Wuppertal in 1974.

[8] In a more explicit treatment of these facts one would have to append a further equation to the XP position which shows that XP controls the gap created by topicalisation (e.g. in "was er hat gegessen (gap)?" as compared with "er hat gegessen ein Kuchen" (both examples in x+1 interlanguage). For the purpose of this paper I will not include this aspect of the description.

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

or "PART") can be described as a gradual lexical acquisition process which is based on a number of alterations of the existing c-structure rule as shown in (R3). One alteration concerns the introduction of VP as a constituent, which is necessary to account for a range of phenomena in German, as we will see below. The other alteration is concerned with the position of the verb. VP rewrites alternatively into the structure known from R1, or as V-COMP, and the latter constituent rewrites as (NP_{obj1}) (NP_{obj2}) V. This ensures that V will only occur in second position unless licensed by a V that takes V-COMP.

(R3) $S \rightarrow NP_{subj}\ VP$

$VP \rightarrow V \begin{Bmatrix} (NP_{obj1})(NP_{obj2}) \\ \text{V-COMP} \end{Bmatrix}$

$\text{V-COMP} \rightarrow (NP_{obj1})(NP_{obj2})\ V$

Apart from this change in c-structure rules, I assume that the learner gradually re-analyses the verbs of his/her interlanguage, by analysing AUX and V as two separate entries and by adding the feature AUX to the lexical features of V.

To achieve the split verb effect, the newly created auxiliaries and modals are treated as main verbs (with the feature AUX that takes the value "+"), which take VP complements (as in Kaplan and Bresnan, 1982; Netter, 1987). Let us take sentence (3) as an example:

(3) er *hat* ein Bier *getrunken*
he has a beer drunk
"he has drunk/drank a beer"

The simplified lexical entries for the verbs in (3) are as shown in Figure 2. This set of entries and rules, etc., ensures two things which are of relevance here.

(1) A particular (at this stage not necessarily the correct) morphological form of the main verb is used with the auxiliary to express the intended grammatical function. This is achieved by functional well-formedness conditions which ensure that functional annotations match across related constituents. In this case it is the value PAST in (PARTICIPLE) = PAST and (V-Comp PARTICIPLE) = PAST which allows a unification of these two functions and thus legitimates these two constituents in this particular sentence.

getrunken: V,	PRED = "trinken (SUBJ)(OBJ)"
	PARTICIPLE = PAST
	INF = ge
	OBJ CASE = accusative
	AUX = –
hat: V,	PRED = "haben, V-COMP (SUBJ)"
	TENSE = PAST
	AUX = +
	NUM = SG
	PERSON = 3RD
	V-COMP PARTICIPLE = PAST
	V-COMP INF = $_c$ ge

Figure 2. Simplified lexical entries for the verbs in (3).

(2) The second point is that the c-structure rules, in conjunction with the unification processes mentioned under (1), ensure that the two verbs appear in a split position and that only the lexical verb can appear in final position. Figure 2 illustrates why, according to the rule system developed above, only lexical verbs can occur in final position: the PRED value for "hat" contains V-COMP and SUBJ, while that of "getrunken" contains SUBJ and OBJ. The SUBJ of "getrunken" needs to be unified with the SUBJ of "hat" since it is not directly linked to any argument. Because of these differences in the lexical entries of the verbs, and the way they interact with c-structure, "hat" cannot be inserted under that V that is dominated by V-COMP, i.e. "hat" in final position is excluded.

In essence, this means that the positioning of verbs is controlled by the unification of the feature PARTICIPLE. This grammatical system can account for what seems to be an unsystematic behaviour on the part of the learner, who applies PART (= split verb) only in a certain percentage of all contexts. In other words, the rule is applied with some verbs but not with others. This phenomenon can be accounted for by the fact that some AUX-V combinations can continue to be analysed as one lexical entry while others are not. What determines this variational phenomenon, then, is whether the verbs in question have been analysed as single lexical entries and whether the feature AUX has been appended and annotated correctly.[9]

To account for German Subject-Verb INVERSION, c-structure has to be modified further. The modifications suggested here are adaptations from Kaplan and Bresnan's (1982) and Pinker's (1984) treatment of inversion in English, which assumes that there is an optional Verb to the left of S as illustrated below:

$\bar{S} \rightarrow (V)\ S$

Pinker adds the constraining equation $ROOT =_c +$ to the verb position in this rule to ensure that inversion only applies to matrix sentences (i.e. the feature ROOT is constrained to be + in matrix and – in embedded clauses). This distinction is also relevant

[9] Note that while these rules account for the facts observed in L2 acquisition, the reality of German word order is far more complex.

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

to the analysis of Standard German, where INVERSION is blocked in embedded clauses. Pinker (1984) further adds the constraining equation *SENT MOOD* $=_c$ *INV* to the verb position in order to be able to allow the rule to constrain INVERSION *lexically* in elements which can occur in topicalised position (cf. (R2) above). The resulting rule is given in (R4).

(R4) $\bar{S} \rightarrow (V) \qquad S$

$$\left\{ \begin{array}{l} \text{ROOT} =_c + \\ \text{SENT MOOD} =_c \text{Inv} \end{array} \right\}$$

In essence, the same rule also operates in German. It is the interaction of (R2) and (R4) which creates the correct word order: a lexical entry for adverbs such as "selten" (seldom) or a lexical redundancy rule for wh-words and prepositional phrases etc. ensures that the filling of the focus position creates the information "sentence MOOD = inv."[10] This information then feeds into the equation in (R4) which licences a verb in a position left of NP_{subj}. In other words, grammatical information is created through the processing of one constituent, and that information is being utilised during the processing of another constituent.

On the basis of the picture that has emerged so far, word order in subordinate clauses has been treated in exactly the same way as in matrix clauses. However, at stage x+5 the learner starts to distinguish between matrix and subordinate clauses. This is evidenced by the final positioning of verbs: $[XV_{INF=-}\ V_{INF=+}]_S$, where *INF* = − refers to verbs not marked for person or number and *INF* = + refers to verbs which are marked for those features. At the same time, INVERSION disappears (sometimes gradually) from subordinate clauses. To account for these facts, the feature ROOT has to be introduced, i.e. the distinction between matrix and subordinate clause. This has the effect that INVERSION would be ruled out in subordinate clauses, assuming that Pinker's constraining equation *ROOT* $=c$ + is appended to V in (R4).[11]

If one accepts this rough proposal for the treatment of German interlanguage word order, then the next point in my argument will be to show that the hierarchy of processing resources developed in the previous section can be incorporated into this description and that the combination of the two elements accounts for the orders of acquisition discussed above.

My account of German word order development started with phase 2 of the acquisition process which is characterised by a strict SVO word order. Since grammatical functions are assigned at the level of c-structure, a strict canonical order does not involve any feature unification and therefore corresponds to level 2 of the hierarchy of processing resources developed in the first section above. In other words, the LFG account of this structure positions it correctly in the hierarchy of processability and its actual phase of acquisition.

German stage 3 syntax was accounted for by a modification of the stage 2 c-structure rule. This modified rule also does not involve any exchange of grammatical information. From a processing point of view, the difference between SVO and ADV is the following. While SVO is a completely linear structure with NP_{subj} in a predictable and invariable position, there is a degree of non-linearity in ADV where the sequence of constituents may deviate somewhat from a strictly canonical sequence. In the latter case, the canonical sequence starts after the topicalised phrase. To achieve this process the learner can utilise the general cognitive mechanism of saliency which allows the developing grammatical encoder to process a position that is external to the canonical sequence. However, in contrast to structures which are acquired later, in the structure ADV-SVO, grammatical functions can be read directly off c-structure and no cross-constituent unification is required.

For the German split-verb construction to occur, the PARTICIPLE value of the main verb and that of V-COMP in the auxiliary entry have to be unified. This exchange of information occurs across constituent boundaries and the matching of the PARTICIPLE value occurs in the VP procedure. Therefore PART requires processing resources which are located higher on the implicational hierarchy than SVO and ADV, i.e. level 4 of the processability hierarchy. Note that the only non-canonical position involved in this process is one that is perceptually salient, namely the lexical verb in final position.

Subject-verb inversion (INV), then, involves a process that depends on the unification of the feature SENT MOOD across XP and V in the S-procedure. Since the S-procedure is hypothesised to become available at level 5 in the processability hierarchy, INV is positioned at that level. Note that this process cannot rely on the saliency principle.

The German structure V-Final is one of the features distinguishing embedded from matrix clauses in the target language. In the above LFG description this structure is accounted for by the introduction of the feature ROOT into the IL grammar and a separate c-structure for embedded sentences. This description accounts for the fact that in the processability hierarchy, features of embedded clauses which

[10] For further detail see Pienemann (in press).

[11] In Pienemann (in press) I develop a more differentiated approach to this problem that accounts for the gradual nature of the acquisition of verb final. However, the details of that argument are not crucial here.

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

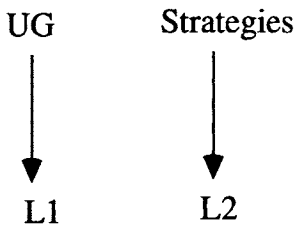


Figure 3. Clahsen and Muysken's (1986) view of the relationship of explanandum and explanans.

distinguish those from matrix clauses are acquired after word order constraints in the matrix clause have been acquired. In other words, the above account of V-Final is in line with the processability hierarchy which predicts that V-Final occurs at level 6.

In summary, I characterised the sequence in the acquisition of German L2 word order in terms of unification of lexical features in a modified LFG framework. I showed that the unification processes involved follow the sequence predicted by Processability Theory. Hence the theory accounts for the observed sequence.

Comparing L1 and L2 acquisition

A number of rationalist approaches to SLA have in common the assumption of *fundamental differences* in first and second language acquisition (Felix, 1984; Clahsen, 1988, 1990; Meisel, 1991; cf. also Bley-Vroman, 1990). They assume that L1 learners have access to UG and that L2 learners[12] do not. To account for L2 acquisition, these authors therefore make recourse to a processing-oriented alternative to UG. This position has become known as the Fundamental Difference Hypothesis (Bley-Vroman, 1990). Clahsen and Meisel are the scholars who have produced the most explicit accounts of explanations of L2 acquisition which are conceived as learning and processing strategies. Clahsen and Muysken (1986) view the relationship of explanandum and explanans roughly as in Figure 3.

In this section I will argue for a different relationship. In particular I will argue that the notion of Universal Grammar (UG) and language processing do not form a dichotomy in the context of explaining differences between L1 and L2 acquisition. Above, I made the case that the fundamental principles of language processing apply to native and non-native language use. I therefore argue that the architecture

[12] I use the term "L2 learner" here to keep the text readable. However it should be noted that most authors base their hypotheses on the critical age of around 6 years. I will return to this point later in the text.

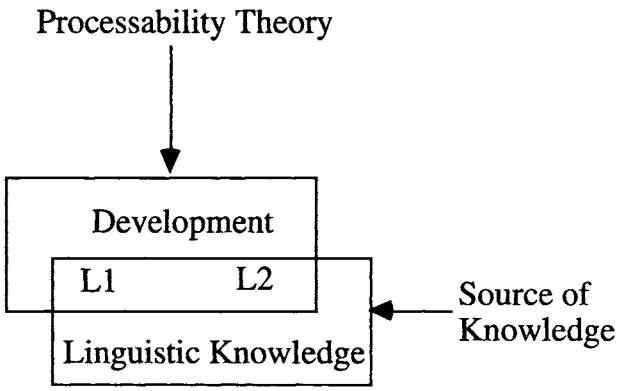


Figure 4. Author's view of the relationship of explanandum and explanans.

of human language processing will have a bearing on any type of language acquisition. In fact, the two explanatory devices, UG and strategies, are on quite a different scale and address different aspects of the acquisition process.

UG has been productive mostly as a property theory, addressing the issue of the origin of linguistic knowledge (i.e. the "logical problem") and has been far less successful in accounting for the "developmental problem", for which a transition theory is needed. I made the point above that Processability Theory is designed exclusively to address the developmental problem, and I will show that it accounts for development, not only in the L2 context but also for L1 and that it can interact with other theory modules which do address the logical problem. I therefore view the relationship of explanandum and explanans roughly as in Figure 4 where the processability components address the developmental problem while linguistic knowledge is created by a source that I leave unspecified for the time being. Some researchers would see UG in its place. This view allows the Fundamental Difference Hypothesis to be maintained without attributing the L1–L2 differences to a processing factor that applies equally to both types of acquisition.

Before I apply Processability Theory to L1 acquisition, let us first have a brief look at an account of the key descriptive facts relating to L1 acquisition. Long (1988, 1990) presents extensive evidence in support of the view that there are marked differences between L1 and L2 acquisition in *ultimate attainment*. He demonstrates that age of onset "... is a robust predictor of their [the L2 learners', MP] long-term success ..." (Long, 1988, 197). Remarkable differences between L1 and L2 acquisition also exist in the developmental schedule. In his 1982 study, Clahsen found the following developmental pattern in the acquisition of German as a first language:

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

L1 sequence

(1) variable word order,
(2) SOV,
(3) V-2nd and SV-agreement marking,
(4) subordinate clauses (without any mistakes in the positioning of the verb).

Quite clearly, this developmental pattern differs markedly from the one observed in the acquisition of German as a second language which was discussed in the previous section.

It is important to note that the differences between L1 and L2 go beyond that of the developmental path to include the following phenomena:

- The virtually error-free learning of the position of the verb in subordinate clauses. Clahsen observed that as soon as the child uses complementisers the position of verbal elements in subordinate clauses is completely in line with the structure of the adult language.
- The interlocking of different rules in the acquisitional process. Clahsen found that in German child language development SV-agreement is acquired exactly at the same point in time as the V-2nd position. This is not the case in the acquisition of German as L2. The two rules were not found to interlock closely. I found the same in a longitudinal study of the acquisition of child GSL (Pienemann, 1981).

The issue is how to explain these differences. In Clahsen and Muysken's (1986) view, all of these differences are due to the fact that children have access to Universal Grammar while adults do not. Instead, adults use inductive learning strategies, which enable them to reconstruct the canonical sentence schema of the target language. This implies that L1 learners are able to discover the underlying word order for German, whereas L2 learners simply infer the canonical sentence schema of German which reflects a syntactic feature of the surface structure of the target language.

Clahsen (1984) and Clahsen and Muysken (1986) account for the developmental path of L2 acquisition using a set of processing strategies. The latter have been subject to serious criticism of which I will mention only three key issues. White (1989) points out that Bever's (1970) strategies were designed for language comprehension. However, Clahsen applied these to production data. White further demonstrates that the assumed strategies do not contain any information or mechanisms that allow for their further development. This is known as the property of extensibility that a theory of learnability must entail (Pinker, 1984). In addition, Clahsen's strategies were conceptualised as processing constraints on transformations as described in "transformational grammar". However, it is now accepted that transformations are psychologically implausible concepts (Altmann, 1990; Horrocks, 1987; Levelt, 1989; cf. also Ingram, 1971). As a result of this, it is illogical to assume that processing constraints can operate on linguistic structures which have no psychological plausibility. In short, the processing explanation invoked for L2 acquisition by Clahsen in support of the Fundamental Difference Hypothesis is not viable. In the previous section I demonstrated that it can be replaced by Processability Theory in the L2 context.

Clahsen and Muysken (1986) argue that the interlocking of different rules in the acquisitional process can be attributed to the setting of the INFL parameter, which is responsible for the distinction between finite and non-finite verbal elements, the crucial distinction that is the prerequisite for both rules. They further claim that the grammatical systems produced by L2 learners are not possible grammars, and in particular that the rules necessary to derive the above L2 patterns from an underlying SVO order are not possible in terms of UG. This claim has inspired du Plessis, Solin, Travis, and White (1987) and Schwartz (1988) to propose a reanalysis of the above L2 sequences which is indeed based on an underlying SVO order.

However, Meisel (1991, 237) points out that ". . . du Plessis et al (1987), in order to be able to write a grammar for L2 learners compatible with UG, have to postulate the existence of two more parameters . . . ." He notes that Schwartz's analysis implies a similar requirement and that in addition, both proposals rely on a poorly defined process of restructuring. Meisel quotes Chomsky (1981) in showing that the ad hoc nature of these proposals runs counter to the fundamental nature of parameters to form clusters of a number of seemingly unrelated grammatical phenomena. In view of these shortcomings he concludes that the above counter-proposals remain unconvincing.

Let us recapitulate. Clahsen and Muysken (1986) and Meisel (1991, 1997) present evidence in favour of access to UG for L1 acquisition and evidence of limited access for L2 acquisition. I have shown that the developmental problem in SLA can be explained by processability. It is therefore quite logical to ask if Processability Theory also accounts for the developmental problem in L1 acquisition.

To test this hypothesis it has to be demonstrated that the above L1 sequence is positioned within the constraints defined by Processability Theory. Formally, this is achieved by showing that the L1

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

sequence is predictable by the hierarchy of processing resources on which Processability Theory is based.

Similarly to SVO structures in L2 acquisition the initial word order hypothesis in L1 acquisition (i.e. SOV – after variable word order) can be accounted for simply by a c-structure rule along the lines of (R5). Since grammatical functions can be read off c-structure canonical orders and does not involve any transfer of grammatical information, the SOV order is positioned at the lowest level in the processability hierarchy.

(R5) $S \rightarrow NP_{subj}\ VP$
$VP \rightarrow (NP_{obj1})(NP_{obj2})\ V\ (V)$

(R6) $\bar{S} \rightarrow (XP)\ \left\{ \begin{array}{l} wh =_c + \\ adv =_c + \\ N =_c + \\ SENT\ MOOD = Inv \end{array} \right\}\ S$

(R7) $\bar{S} \rightarrow (V)\ \left\{ \begin{array}{l} ROOT =_c + \\ SENT\ MOOD =_c Inv \end{array} \right\}\ S$

The Verb-2nd phenomenon can be produced by (R6) and (R7) in a way similar to German and English INVERSION. Note that the constraint equation ROOT $=_c$ + is not appended to V at this stage. For the V-2nd position to be produced, the grammatical information SENT MOOD has to be exchanged by two constituents (XP and V). This places V-2nd at the same level in the processability hierarchy as INVERSION (in the L2 sequence) and SV-agreement.

One reviewer commented that one might conclude from this that Processability Theory entails the prediction that the above structures are acquired simultaneously. However, it would be incorrect to derive such a prediction from the theory, since it does not predict that whatever can be processed will indeed be acquired. Instead, the theory predicts that what cannot be processed will not be acquired. In other words, processability acts as a constraint on development and therefore does not entail the prediction that the above structures are acquired simultaneously.[13]

The sentence-final position of the verb in subordinate clauses can be accounted for by appending the constraint equation $ROOT =_c +$ to V in (R-c). Since the feature $\pm ROOT$ emerges at level 6 of the processability hierarchy, the final stage of the L1 sequence is also in line with Processability Theory.

In other words, SOV, V2nd and V-Final (as well as SV-agreement) do indeed fall within the constraints of Processability Theory, and the L1 sequence is explained by the same hierarchy of processability as the L2 sequence. I hasten to add that this does not imply that the two processes are one and the same thing. To start with, the routes of acquisition are different. The reader will notice that the rule PART is absent from the L1 sequence. To explain why this is structurally possible one has to consider the effect of the rules (R5)–(R7): starting (developmentally) with an SOV c-structure, these three rules have the same effect as the combined application of PART and INVERSION on the basis of an SVO c-structure. Since in (R5) the verb is in final position, and (R6) jointly with (R7) permit the finite verb to appear in second position, the "split verb" position is also permitted.

The rule ADV is also absent from the L1 sequence. This means that L1 learners do not produce the ungrammatical structure X S V Y which is typical for L2 learners. The processability hierarchy allows for ADV to emerge at level 3. However, it does not predict that it will necessarily develop.

Table 1 gives an overview of this comparison of grammatical development in the acquisition of German as a second and as a first language which shows at a glance that both developmental paths fall within the confines of Processability Theory. In other words, there are no differences in the temporal order in which processing resources are activated. All grammars are processable at the time they develop, and each grammar builds upon the processing resources acquired at the previous stages in a cumulative fashion. However, the L1 learner achieves this in two key "moves", SOV and V2nd (with SV agreement) while the L2 learner takes five "moves", most of which introduce ungrammatical structures which have to be modified in later moves.

We now find ourselves in a situation where strikingly different developmental routes in L1 and L2 acquisition have been accounted for within one and the same hierarchy of processing resources. Therefore the formula "UG for L1 and processing factors for L2" no longer holds. And one has to ask oneself what causes the apparent differences between L1 and L2 that exist despite the common basis in language processing.

Developmental dynamics and generative entrenchment

My basic thesis is that different outcomes and developmental paths in language development are, at least

[13] In contrast, parametrization is a property theory which addresses epistemological issues in language acquisition. It specifies which bits of linguistic knowledge the learner has access to through UG. Since the latter theory links the two phenomena in question to a single parameter, the setting of this parameter necessarily includes the prediction of the simultaneous acquisition of the two structures.

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

Table 1. *Overview of grammatical development in German L1 and L2*

Stage	Exchange of information	Resources	GSL	German L1
6	within subor. clause	+/- ROOT	V-End	V-End (no errors)
5	inter-phrasal	WO rules S-Procedure	INV ±agr	V2nd +agr
4	phrasal	WO rules VP-Procedure saliency	PART	–
3	none	lexical categories saliency	ADV	–
				b SOV
2	none	lex. categories	SVO	a variable word order
1	none	lexical entries	words	words

partly, due to different developmental dynamics, caused by differences in the initial hypotheses and that the process of development can be fundamentally similar, with respect to language processing, despite fundamentally different outcomes and different developmental paths.

The basic mechanism behind developmental dynamics is the principle that developmentally early decisions bias the further development of the interlanguage system. This percolation of structural properties in developmental processes is known in biology and philosophy and has been termed "generative entrenchment" by Wimsatt (1986, 1991).

The concept of generative entrenchment is exemplified, for instance, by the embryonic development of animals where sections of the fertilised egg take on more and more specialised structures (e.g. Gehring, 1985; Coen and Carpenter, 1992; Wolpert, 1992). The segmentation of the body plan occurs very early in these processes for all animals. In other words, the position of head, limbs, etc. is determined very early. These structural features are maintained throughout the developmental process, and they do not have to be decided on every time a refinement of parts of the structure is made. One can say that these features are "developmentally entrenched."

We also know that incorrect information on the positioning of segments can have serious consequences for the ultimate shape of the organism (Gehring, 1985; Coen and Carpenter, 1992; Wolpert, 1992). This sometimes unfortunate phenomenon illustrates the concept of the *depth* of generative entrenchment. The earlier a decision is made in

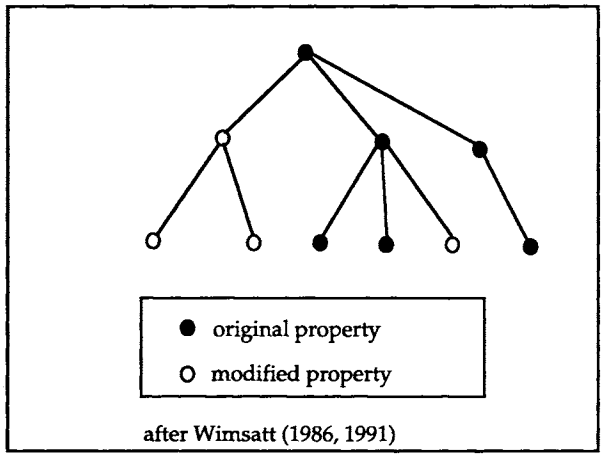


Figure 5. Illustration of structural development as generative entrenchment.

structural development, the more far-reaching the consequences for the ultimate stage in structural development.

Figure 5 illustrates how development can be understood as a generative process where structures increase in complexity, starting with a minimal number of structural properties to which other properties are added throughout development. Figure 5 displays structural options that exist in development of structure as a tree diagram where each node represents a point at which more specialised structures can develop. The top node contains the initial structural information. The tree allows for different developmental paths. However, once a decision has been made and a new structure has been added, it is very costly, if not impossible, for the developmental process to move to a different developmental path. In effect, changing the developmental path would mean that all developmental steps up to the node that gives access to the alternative path would have to be cancelled. As a result, a great deal of structural information would be lost in such a move. Many physical processes of development are indeed irreversible, as the example of developmentally misformed organisms shows.

The key explanatory point that can be derived from the concept of generative entrenchment for language acquisition is that a massive computational saving can be made if structural decisions do not have to be revised in the developmental process every time a structural change occurs. In this model initial structural features propagate in the developing system and thus determine the ultimate structure without being invoked again and again. The basic "body plan" stays the same. In other words, a computational saving is made by laying structures down and keeping them. The alternative would be a

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

developing system in which all processes of structural refinement have to be orchestrated globally for every developmental step, and this would require far more computational resources than the preservation of structures once they have developed.

The aspect of computational saving inherent in generative entrenchment is captured by Wimsatt (1986, 1991) in his "developmental lock" which is based on Herbert Simon's (1962) classic paper, "The Architecture of Complexity," where he demonstrates that solutions of complex problems can be found more effectively by using the heuristic of factorising sub-problems which are solved independently, and the solutions to sub-problems are strung together to produce the solution to the overall problem.

Wimsatt's developmental lock is an idealised set of complex problems. The lock consists of ten wheels with ten positions each, very much like an extended version of a combination lock. Obviously, the total number of possible combinations on this lock is 10^{10}, which, according to Wimsatt, requires 10^9 trials to find the correct combination. In this form of unconstrained hypothesis testing, the lock is referred to as a "complex lock". In the developmental lock, Wimsatt constrains hypothesis testing by allowing the problem solver to factorise the combination problem. Rather than having to get all ten digits right before the combination can be subjected to an empirical test, each wheel can be tested individually in a left-right sequence. This additional ability to test the wheels individually from left to right mimics the cumulative nature of developmental processes: old solutions can be kept.

The computational advantage of factorising the complex problem is remarkable. Only fifty trials are necessary to find the solution to the developmental lock problem, providing a strict left-right sequence is followed. In other words, in this approach later decisions depend on earlier decisions. If an error is made earlier it will be very costly to recover from it, because all intervening solutions will be lost.

Summing up, the assumption that underlies the notion of generative entrenchment is that, in the course of development, structures are preserved and further refined. The dynamics of this process have a constraining effect on development: because of the preservation constraint, early decisions exclude whole branches on the "developmental tree".[14] In other words, generative entrenchment can be understood as a constraint on development that derives from the dynamics of development itself. In this way it complements the notion of processability which also acts as a constraint on language development.

Hence, the concept of generative entrenchment is a general logical-mathematical set of constraints that applies to any continuous developmental process in which structures diversify developmentally. It has been applied to many cases of biological development in ontogenesis and phylogenesis as well as population development. It is also applicable to the development of physical phenomena. Since language acquisition is another case in which structure diversifies developmental from an initial state, the dynamics of generative entrenchment apply. An anonymous reviewer objected that language acquisition is different from cases of biological and physical development, because the role of input and environment is different in the case of language development. However, this objection missed the point. Generative entrenchment models nothing but the dynamics of cumulative developmental processes, no matter where they occur. All of the non-language cases of development are also placed in an environment, be it the distribution of matter in space or the location of a body segment of the fruit fly embryo in an acidic environment. Generative entrenchment should be seen on the same scale as the logical problem where one looks at the logic of developmental dynamics.

The continuity assumption that underlies the notion of generative entrenchment can be verified in the two developmental paths under discussion. Table 2 reveals a great deal of continuity in both developmental paths. The initial word order hypothesis of the L2 learner is maintained and refined in four developmental steps until it is finally modified to reflect the structure of the TL. It is evident from distributional analyses of large bodies of data (cf. Clahsen, Meisel, and Pienemann, 1983) that the SVO pattern continues to be produced in tandem with INVERSION and V-End. Looking at the L1 perspective, a similar picture emerges: the initial SOV pattern is maintained and refined throughout the developmental process.

In other words, pattern conservation is evident in the two developmental paths, and this gives a first indication of the initial hypothesis being a determining factor for the later course of development. However, there is also direct evidence of this. The need for the separate rule PART arises only if the initial syntactic hypothesis is SVO. In order to modify the initial SVO order to match German main clause patterns, two developmental "moves" are needed: one that ensures that V_i is in final position and one that constrains V_f into second position. INV on its own would not be the solution. If INV were applied without PART, ungrammatical structures

[14] There is ample evidence for this in biological development (cf. Gehring, 1985; Coen and Carpenter, 1992; Wolpert, 1992).

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

Table 2. *Generative entrenchment of linguistic structures*

	L2	L1
subordinate clause	comp SOV [comp ***S V O***]	comp ***SOV***
use of S-procedure for storage across constituents in S	±agr X Vf S O Vi [(X) ***S V O***]	+agr X Vf ***S O V***i
use of VP-procedure for storage across constituents in VP	X ***S Vf O V***i	–
use of saliency principle to relax canonical order constraint	X ***S V O***	–
word order constrained into canonical order	***S V O***	***S O V***

like $X\ V_f\ S\ V_i\ Y$ would result.[15] In other words, both PART and INV are necessitated by the initial SVO word order and the initial SVO hypothesis is propagated in the development of these rules.

Based on an SOV order, V2nd produces the same effect as PART and INV combined, namely the correct target structure $X\ V_f\ S\ Y\ V_i$. This means that the development from SOV to V2nd renders the acquisition of PART and the two developmental steps PART and INV superfluous.

A further developmental saving is made in the acquisition of the verb final position in subordinate clauses. In L1 acquisition, where the initial fixed word order is SOV, the position of the verb does not have to change when the distinction between main and subordinated clauses is acquired. All that has to happen is that V2 is blocked in subordinate clauses. This allows verb-final positions to be acquired with less effort than in the L2 sequence, where every complementiser has to be marked for the feature ±ROOT until a lexical redundancy rule is formed for subordinate clauses. And yet again this developmental twist is brought about by the moves that followed from the initial structural hypothesis. This explains why in L1 acquisition the verb-final position is acquired virtually without errors, while in L2 acquisition the development of the rule can be traced along the path of an increasing set of complementisers with the effect of an error-ridden acquisition process.

[15] This structure has not been reported in GSL acquisition studies.

In summary, the initial hypothesis for word order in L1 acquisition renders the acquisition of PART and V-End as separate c-structure rules superfluous and results in a grammatically more correct output during development. Hence, the L1 initial hypothesis is far more economical, and so is the development that follows from it.

One might, of course hypothesise that L2 learners fundamentally restructure the developing grammar and switch to the assumption that German is an SOV language (e.g. Tomaselli and Schwartz, 1990). However, it is hard to see, in the absence of negative evidence, on the basis of which structural evidence the learner would arrive at such a conclusion, since the SVO structures they produce form a subset of the input data. But even if the learner were able to arrive at such a conclusion by some yet unknown means, a total restructuring of the interlanguage would be computationally very costly since all rules developed until the restructuring would have to be abandoned to start a new developmental path. Neither SVO nor PART nor INV is compatible with the SOV assumption. In this way, the developmental path known in L1 acquisition becomes inaccessible once the initial syntactic hypothesis has been formed, and the development that follows is determined by the initial hypothesis, because the structure of the target language can be mapped only by the set of rules ADV, PART, INV and V-END, and these rules emerge in this sequence because of their processability.

From the above discussion, we can summarise the following points:

- there are two distinct developmental paths;
- each of them falls within the constraints defined by Processability Theory;
- one of the paths is superior to the other;
- each is determined by the initial hypothesis.

One can conclude from this that one of the main reasons for L1–L2 differences is the superior initial hypothesis of the L1 learner which propagates through the entire developmental process. This raises the obvious question as to what it is that causes the initial hypothesis. However, this question goes well beyond the developmental problem which I set out to address with Processability Theory. The source of the initial hypothesis is an epistemological issue which requires not a transition theory (such as Processability Theory) but a property theory (cf. Gregg, 1996).

Final remarks

The purpose of this paper was twofold: (1) to give an overview of the theory of processability in language acquisition, and (2) to apply this theory to a compar-

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

ison of developmental dynamics in L1 and L2 acquisition.

I demonstrated that the hypothesised processability hierarchy accounts for the developmental routes found in both types of acquisition even though these routes are fundamentally different. Processability Theory affords a new perspective on the L1–L2 comparison in which the architecture of human language processing has a bearing on both types of acquisition. L1 and L2 learners produce different initial hypotheses, and the structural effect of these hypotheses biases the subsequent development through the propagation of earlier structural features.

These developmental dynamics account for major aspects of the L1–L2 developmental difference. They do not, of course, settle the issue of which epistemological source L2 learners do have access to and how procedural skills are acquired. Those issues need to be tackled by different theory modules. What Processability Theory does address is the developmental problem in language acquisition, and there is a wide range of further contexts, such as typological differences and learner variation, in which the theory is productive (cf. Pienemann, in press).

References

Altmann, G. T. M. (1990). Cognitive models of speech processing: An introduction. In Altmann, G. T. M. (ed.), *Cognitive models of speech processing. Psycholinguistic and computational perspectives*, pp. 1–23. Cambridge, MA: MIT Press.

Baddeley, A. (1990). *Human memory: theory and practice*. Hillsdale: Erlbaum.

Bates, E., MacWhinney, B., & Smith, Stan (1982). Pragmatics and syntax in psycholinguistic research. In S. W. Felix & H. Wode (eds.), *Language development at the cossroads: Papers from the interdisciplinary conference on language acquisition at Passau*, pp. 11–30. Tübingen: Gunter Narr Verlag.

Berwick, R. & Weinberg, A. (1984). *The grammatical basis of linguistic performance: language use and language acquisition*. Cambridge, MA: MIT Press.

Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. Hayes (ed.), *Cognition and the development of language*. New York: Wiley.

Bley-Vroman, R. (1990). The logical problem of second language learning. *Linguistic Analysis*, 20, 3–49.

Bock, M. (1978). *Wort-, Satz- und Textverarbeitung*. Stuttgart: Kohlhammer.

Braine, M. D. S. (1976). Children's first word combinations. *Monographs of the Society for Research in Child Development*, 41, Serial #164.

Braine, M. D. S., & Wells, R. S. (1978). Case-like categories in children: The actor and more related categories. *Cognitive Psychology*, 10, 100–22.

Bresnan, J. (ed.) (1982). *The mental representation of grammatical relations*. Cambridge, MA: MIT Press.

Broadbent, D. E. (1975). The magic number seven after fifteen years. In A. Kennedy & A. Wilkes (eds.), *Studies in long term memory*. London: Wiley.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Floris.

Clahsen, H. (1980). Psycholinguistic aspects of L2 acquisition. In S. W. Felix (ed.), *Second language development*, pp. 57–79. Tübingen: Narr.

Clahsen, H. (1982). *Spracherwerb in der Kindheit. Eine Untersuchung zur Entwickung der Syntax bei Kleinkindern*. Tübingen: Narr.

Clahsen, H. (1984). The acquisition of German word order: A test case for cognitive approaches to L2 development, pp. 219–42. In R. Anderson (ed.), *Second languages*. Rowley, MA: Newbury House.

Clahsen, H. (1988). Parameterised grammatical theory and language acquisition: A study of the acquisition of verb placement and inflection by children and adults. In S. Flynn & W. O'Neill (eds.), *Linguistic Theory in Second Language Acquisition*, pp. 47–75. Dordrecht: Kluwer.

Clahsen, H. (1990). The comparative study of first and second language development. *Studies in Second Language Acquisition*, 12, 135–53.

Clahsen, H. (1992). Learnability theory and the problem of development in language acquisition. In J. Weissenborn, H. Goodluck & T. Roeper (eds.), *Theoretical issues in language acquisition: Continuity and change*, pp. 53–76. Hillsdale, NJ: Lawrence Erlbaum.

Clahsen, H., Meisel, J. & Pienemann, M. (1983). *Deutsch als Zweitsprache: Der Spracherwerb ausländischer Arbeiter*. Tübingen: Narr.

Clahsen H., & Muysken, P. (1986). The availability of universal grammar to adult and child learners: A study of the acquisition of German word order. *Second Language Research*, 5, 93–119.

Coen, E., & Carpenter, R. (1992). The power behind the flower. *New Scientist*, 25, April, 26–9.

Cooper, W. E., & Zurif, E. B. (1983). Aphasia: information-processing in language production and reception. In B. Butterworth (ed.), *Language Production*, vol. 2. London: Academic Press.

De Bot, K. (1992). A bilingual production model: Levelt's speaking model adapted. *Applied Linguistics*, 13, 1–24.

Du Plessis, J., Solin, D., Travis, J., & White, L. (1987). UG or not UG, that is the question: a reply to Clahsen and Muysken. *Second Language Research*, 3, 56–75.

Engelkamp, J. (1974). *Psycholinguistik*. München: UTB.

Engelkamp, J. & Zimmer, H. D. (1983). *Dynamic aspects of language processing*. Heidelberg: Springer Verlag.

Felix, S. W. (1984). Maturational aspects of universal grammar. In A. Davies, C. Criper, & A. Howatt (eds.), *Interlanguage*. Edinburgh: Edinburgh University Press.

Felix, S. W. (1991). Accessibility of UG in second language acquisition. In L. Eubank (ed.), *Point–counterpoint: Universal Grammar in the second language*, pp. 89–104. Amsterdam: John Benjamins.

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

Garman, M. (1990). *Psycholinguistics*. Cambridge: Cambridge University Press.

Garrett, M. F. (1976). Syntactic process in sentence production. In R. Wales & E. Walker (eds.), *New approaches to language mechanisms*, pp. 231–56. Amsterdam: North Holland.

Garrett, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (ed.), *Language production*: Vol. 1, *Speech and Talk*, pp. 177–220. London: Academic Press.

Garrett, M. F. (1982). Production of speech: observations from normal and pathological language use. In A. W. Ellis (ed.), *Normality and pathology in cognitive functions*, pp. 19–76. London: Academic Press.

Gehring, W. J. (1985). The molecular basis of development. *Scientific American*, 253 (4), 137–46.

Gregg, K. (1996). The logical and the developmental problems of second language acquisition. In W. C. Ritchie, *Handbook of second language acquisition*, pp. 49–81. San Diego: Academic Press.

Horrocks, G. (1987). *Generative Grammar*. London and New York: Longman.

Ingram, E. (1971). A further note on the relationship between psychological and linguistic theories. *IRAL*, 9, (4), 335–46.

Jansen, L. (1991). The development of word order in natural and formal German second language acquisition. In *Australian Working Papers in Language Development*, 5, 1–42.

Kaplan, R., & Bresnan, J. (1982). Lexical-Functional Grammar: a formal system for grammatical representation. In J. Bresnan (ed.), *The mental representation of grammatical relations*, pp. 173–281. Cambridge, MA: The MIT Press.

Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science*, 11, 201–58.

Kintsch, W. (1970). *Memory and Cognition*. New York: Wiley.

Kintsch, W. (1974). *The representation of meaning in memory*. Hillsdale: Erlbaum.

Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition*, 14, 41–104.

Levelt, W. J. M. (1989). *Speaking. From intention to articulation*. Cambridge, MA: MIT Press.

Long, M. H. (1988). Second language acquisition as a function of age: research findings and methodological issues. In K. Hyltenstam & Å. Viberg (eds.), *Progression and Regression in language*, pp. 196–221. Cambridge: Cambridge University Press.

Long, M. H. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition*, 12, (3), 251–86.

Marslen-Wilson, W. D., & Tyler L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, 8, 1–71.

Maxwell, J. T., & Kaplan, R. M. (1995). The interface between phrasal and functional constraints. In M. Dalrymple, R. M. Kaplan, J. T. Maxwell and A. Zaenen (eds.), *Formal issues in Lexical-Functional Grammar*. Stanford: CSLI Publications.

Meisel, J. M. (1991). Principles of Universal Grammar and strategies of language use: On some similarities and differences between first and second language acquisition. In L. Eubank (ed.), *Point–counterpoint: Universal Grammar in the second language*, pp. 231–76. Amsterdam: John Benjamins.

Meisel, J. M. (1997). The acquisition of the syntax of negation in French and German: Contrasting first and second language development. *Second Language Research*, 13, 227–63.

Meisel, J. M., Clahsen, H., & Pienemann, M. (1981). On determining developmental stages in natural second language acquisition. *Studies in Second Language Acquisition*, 3, 109– 35.

Murdock, B. B. Jr (1962). The serial position effect in free recall. *Journal of Experimental Psychology*, 62, 618–25.

Netter, K. (1987). Wortstellung und Verbalkomplex im Deutschen. In Schreiber et al. (eds), *Computerlinguistik und Philologische Datenverarbeitung*, pp. 98–114. Hildesheim.

Paradis, M. (1987). *The assessment of bilingual aphasia*. Hillsdale: Erlbaum.

Paradis, M. (1994). Neurolinguistic aspects of implicit and explicit memory: Implications for bilingualism and SLA. In N. C. Ellis (ed.), *Implicit and explicit language learning*, pp. 393–419. London: Academic Press.

Penfield, W., & Roberts, L. (1959). *Speech and brain-mechanisms*. Princeton, N.J.: Princeton University Press.

Pienemann, M. (1980). The second language acquisition of immigrant children. In S. W. Felix (ed.), *Second language development: Trends and issues*, pp. 41–56. Tübingen: Narr.

Pienemann, M. (1981). *Der Zweitspracherwerb ausländischer Arbeiterkinder*. Bonn: Bouvier.

Pienemann, M. (1988). Determining the influence of instruction on L2 speech processing. In G. Kasper (ed.), *AILA Review 5: Classroom Research*, 40–72.

Pienemann, M. (in press). *Language processing and second language acquisition: Processability theory*. MS. Australian National University. Amsterdam: J. Benjamins.

Pinker, S. (1979). Formal models of language learning. *Cognition*, 7, 217–83.

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Schwartz, B. D. (1988). *Testing between UG and problem-solving models of SLA: Developmental sequence data*. MS, University of Geneva.

Shiffrin, R. M., & Schneider, N. (1977). Controlled and automatic human information processing II: attending, and a general theory. *Psychological Review,84*, 127–90.

Simon, H. A. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society*. Reprinted in H. A. Simon (1981), *The science of the artificial*, ch. 7, pp. 192–229. Cambridge, MA: The MIT Press.

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

Sridhar, S. N. (1988). *Cognition and sentence production.* New York: Springer Verlag.

Tomaselli, A., & Schwartz, B. (1990). Analysing the acquisition stages of negation in L2 German: Support for UG in adult SLA. *Second Language Research*, 6 (1), 1–38.

Tulving, E. (1990). Elements of episodic memory. In M. W. Eysenck (ed.), *The Blackwell dictionary of cognitive psychology*, pp. 137–38. Oxford: Basil Blackwell.

Vigliocco, G., Butterworth, B., & Garrett, M. F. (1996). Subject-verb agreement in Spanish and English: Differences in the role of conceptual constraints. *Cognition*, 61, 261–98.

Westmoreland, R. (1983). *L2 German acquisition by instructed adults.* MS. University of Hawaii at Manoa.

Wexler, K. (1982). A principle theory for language acquisiton. In E. Wanner & L. R. Gleitman (eds.), *Language acquisition: The state of the art*, pp. 288–315. Cambridge: Cambridge University Press.

Wexler, K., & Culicover, P. (1980). *Formal principles of language acquisition.* Cambridge, MA: MIT Press.

White, L. (1989). *Universal grammar and second language acquisition.* Amsterdam: Benjamins.

White, L. (1991). Second language competence versus second language performance: UG or processing strategies? In L. Eubank (ed.), *Point–counterpoint: Universal grammar in the second language*, pp. 67–189. Amsterdam: John Benjamins.

Wimsatt, W. C. (1986). Developmental constraints, generative entrenchment and the innate- acquired distinction. In W. Bechtel (ed.), *Integrating scientific disciplines*, pp. 185– 208. Dordrecht: Martinus Nijhoff.

Wimsatt, W. C. (1991). *Generative entrenchment in development and evolution.* MS, Department of Philosophy, University of Chicago.

Wolpert, L. (1992). The shape of things to come. *New Scientist*, 134, 1827 (June), 38–42.

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

Bilingualism: Language and Cognition 1 (1998) 21–22 © 1998 Cambridge University Press

PEER COMMENTARIES

What's in a process? Explaining development in language acquisition

ELLEN BIALYSTOK
Department of Psychology, York University, 4700 Keele Street, Toronto, Ontario, Canada M3J 1P3. ***E-mail:*** *ellenb@yorku.ca*

The tension between considering language acquisition to be a process of learning or a process of development is the visible surface of a theoretical iceberg. The difference between learning and development is neither simple nor self-evident, yet the implications that follow from a commitment to one or the other are considerable. Development is usually construed as an unfolding, the impassive evolution of "being." Children develop physically, cognitively, emotionally almost in spite of themselves; they may not develop to their full potential, and they may require massive environmental support, but even in cases of benign neglect, children will normally become taller, smarter, and more stable. Learning implies a measure of control and responsibility. It depends on experience and on the availability of resources: conceptual, social, intellectual, and so on.

The distinction between formal and functional linguistic theories rests partly on the extent to which theorists consider that language acquisition is a process of learning (functional) or a process of development (formal). Sometimes they coexist. Noting the fundamentally different kinds of experience involved in first and second language acquisition, some researchers have explained them using different presumptions about learning and development (e.g. Clahsen, 1990; Meisel, 1991). Specifically, first language acquisition is left to follow the preordained paths laid down by universal grammar (hence, development), but second language acquisition is characterized as an exercise in problem-solving and active participation by the language learner (hence, learning).

Pienemann's intention is to resolve the problem inherent in proposing two epistemologically different kinds of description to explain two phenomena that have obvious intrinsic similarity. His proposal is a framework he calls "Processability Theory", named for its assertion that constraints derived from human cognitive processing must be included in formal algorithms for language acquisition to account for the grammars that language learners are able to handle at various stages of development. The important claim is that this system can equally explain (and predict) the acquisition of both a first and second language. Furthermore, it obviates the need to distinguish between considering language acquisition as either learning or development, since the two become interconnected.

The theory is an attempt at reconciliation: it acknowledges the role of formal linguistic knowledge (even allowing for its possible innateness), yet emphasizes the need to include a cognitive system that operates on that knowledge in ways commensurate with known processing resources. In this way, it offers a sensible resolution to the apparently intractable dichotomy between learning and development and the implausible wedge it places between first and second language acquisition. Nonetheless, the use of the term "process" and the nature of supporting evidence are problematic.

What does Pienemann mean by "process?" Apart from acknowledging a theoretical lineage to Levelt (1989) and others, he offers no direct definition. He does, however, provide concrete examples of language processing devices: word order rules, syntactic procedures and their specific stores, diacritic features in the lexicon, and so on. Each of these language-specific processes, he argues, needs to be acquired in order to process (learn) the relevant linguistic structure. But how does one distinguish between the process and its attendant structure? Indeed, the so-called processes appear to be slightly abstract formulations of linguistic structures, in other words, knowledge. What is needed here is a different conception of process, one that truly stands outside the specifics of linguistic knowledge. If it could be shown that a hierarchy of cognitive processes influences the order in which linguistic structures can be learned, then that would indeed contribute to the reunification of language acquisition, not only consolidating the acquisition of a first and second language but also integrating language acquisition into broader cognitive domains. The nature of the definition of process adopted in this theory could not meet those goals.

The problem of evidence is a circularity in the argument. The claim is that the production of certain structures is constrained by the availability of the necessary processing resources; the corollary is that structures not supported by such resources will not be produced. The evidence, however, comes only from observed production. Hence, the claim that "Once we can spell out the sequence in which language processing routines develop in the learner, we can delineate those grammars that are processable at different points in development" is fallacious. The "sequence of language processing routines" is derived from the "grammars . . . at different points in development" so cannot be used as evidence for their predicted occurrence.

Finally, a note on generative entrenchment. This is an interesting idea: a structure set up in one stage affects all subsequent developments. The idea is reminiscent of the notion of historical contingency that Stephen Jay Gould (1989) offers as an alternative, indeed, as an antidote, to the more usual logical forms of prediction. Events occur because they follow other events, and altering any incident in the chain would change everything. Language acquisi-

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

tion is exactly like this. Knowledge of language accrues by building on what we already know, not only about language, but also about meanings, communication, structure. Second language acquisition necessarily has different historical contingencies than first language acquisition, not the least of which is that the learner already knows a language. The contingencies may be different, but the process must surely be the same.

Processability theory is a contribution to the tradition of seeking explanations that supersede specific events. We need to be able to look past apparent surface differences to find underlying commonality. Acquisition of language, including first and subsequent systems, rests on both the mechanisms of development that guarantee growth and the processes of cognition that define competence. Pienemann may not have the processes right yet, but the attempt to identify them is a notable start. John Macnamara (1982) ended his important book on language acquisition with the conclusion that the minds of children and adults, despite obvious differences in detail, were the same. He concludes: "That, together with the strategic wisdom of taking two mysteries as one . . . commends the conclusion that they are equivalent" (p. 236).

References

Clahsen, H. (1990). The comparative study of first and second language development. *Studies in Second Language Acquisition*, 12, 135–53.

Gould, S. J. (1989). *Wonderful life*. New York: W.W. Norton.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Macnamara, J. (1982). *Names for things*. Cambridge, MA: MIT Press.

Meisel, J. M. (1991). Principles of universal grammar and strategies of language use: On some similarities and differences between first and second language acquisition. In L. Eubank (ed.), *Point-counterpoint: Universal grammar in the second language*, pp. 231–76. Amsterdam: John Benjamins.

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

Bilingualism: Language and Cognition 1 (1998) 23–24 © 1998 Cambridge University Press

On Processability Theory and second language acquisition

SUSANNE E. CARROLL
Applied Linguistics, Institute for English & American Studies, University of Potsdam, PF 60 15 53, 14415 Potsdam, Germany.
E-mail: *carroll@rz.uni-potsdam.de*

Responding to an invitation to comment on Pienemann's paper leaves me on the horns of a dilemma: I think the approach is novel and shows how future second language acquisition (SLA) research can be done. However, it is difficult to judge from the paper what the exact contribution of the theory is to be.

How can one disagree that Pienemann's approach to explaining SLA, which demands a commitment to an explicit theory of grammar while focusing on the role of processing in describing the developmental stages of L2 speech production, is a good path to follow? Isn't it obvious? No! Processibility Theory research marks a significant departure from the standard literature on SLA, which is only occasionally concerned with describing mental grammars, and almost never concerned with the details of how learners learn them. While the field abounds with interesting, rich, and occasionally important empirical studies, including the ZISA studies in which Processibility Theory is rooted, there are few attempts to write even partial grammars of a learner's interlanguage, and even fewer attempts to answer the question of how those grammars could have been learned from the input the learner has got. Little use is made of contemporary cognitive science research and SLA's contribution to it has been to date nil. Pienemann, in contrast, knows that we need clear models of what grammatical knowledge consists of, combined with equally clear models of processing mechanisms. He ensures that his own work is rooted in a formal theory of grammar, Lexical Functional Grammar (LFG), and Kempen and Hoenkamp's procedural account of speech generation.

It is odd, however, that nothing in Processibility Theory hinges on adherence to LFG. This result is surprising, given the stated focus on learnability. That learnability theory will not tell us how humans learn language is well-known in formal treatments of grammars, and serves as the starting point of debates on feasibility (Chomsky, 1965; Bresnan, 1978; Pinker, 1984; Carroll, 1989). Learnability studies define grammars which are learnable in principle. Feasibility studies define grammars which are learnable in principle under conditions resembling those that humans actually face: limited exposure to stimuli, time constraints on learning, reduced or noisy transmission conditions, limitations on attention and memory, and so on. Since feasibility is not a statement of the factors which determine how a given individual learns some linguistic phenomenon at a given moment, on the basis of given stimuli, provided in a given environment, it stands to reason that defining a feasible grammar is just one step in developing a theory of language learning. Pienemann's paper is not about the nature of grammars. It also makes no independent contribution to spelling out empirically what psychologically real learning conditions are. Therefore, it does not add to either linguistics or psychology's contribution to learnability or feasibility. Moreover, Pienemann grants that Processibility Theory will not account for the contents of the learner's representations. Even allowing for severe constraints on the form of encoding systems provided by Universal Grammar, we still need an account of how speech units are perceived, how categories are formed on the basis of cues from either a system "lower down" in the processing system or a system "higher up", how generalisations are formed, and how exceptions to generalisations are observed and encoded. In my view, this is where most of the action in SLA takes place (see Carroll (submitted)). Finally, the relationship between the hierarchy of processing resources and the stages of development observed in the ZISA study is not a matter of logic. Pienemann, like many others, assumes that one learns "words/lemmas" before one can learn the procedures which combine "words/lemmas" in production. A theory of grammar, however, is not the best source of models for a theory of linguistic memory. What gets stored in linguistic memory may, or may not, correspond to the units grammarians call "words," and what gets put to use in production will consist of all sorts of units, including affixes, morphosyntactic words, phonological words, idioms, and lexicalised "sentence stems" (Pawley & Hodgetts Syder, 1983). We know that learners can isolate long strings, presumably on the basis of their phonetic properties, which means that those strings can be stored and later analysed in the lexicon. Phrasal units need not, therefore, at least in principle, result in production solely from the combination of "words" into phrases by phrasal procedures.

Pienemann's account of development begins only once learners have somehow encoded L2 stimuli in linguistic terms. Is it therefore a reasonable model of bilingual language production? Perhaps not, for nothing in it explains which production procedures will be transferred from the L1 to L2 speech production (or when, or why), nor how learners come to suppress transferred L1 production procedures, which must occur if L2 production is to be fast, accurate and effortless (Green, 1986). It is clear to me that a model designed for explaining production in knowledgeable and fluent monolingual native speakers will not carry over unamended to SLA. Processibility Theory is also not an account of skills development per se, since there is more to be said on that topic than that learners develop new processing procedures and "automatise" them (Pawley & Hodgetts Syder, 1983). What Pienemann has done is to apply a particular formal model of production to a well-established set of developmental facts. I want to emphasise that this effort is anything but trivial. On the contrary, working out the details is difficult, and understanding the

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

limitations of one's own model is hardly straightforward. We can see this work as an effort to take seriously the most sophisticated empirical and theoretical findings of cognitive science research and to test these models against robust data from developing bilinguals. As my commentary should make clear, such models force us to be explicit about language, learning, parsing, memory, and production. Explicitness allows us to formulate questions which could not otherwise be asked. The Processibility Theory exhibits these virtues. Even if the current version leaves many issues unresolved, it will push us to clarify our thinking about how L2 learners perceive, encode, parse and understand L2 speech, and how they produce, monitor, and self-correct it. Somewhere in this equation, if all goes well, we will also find "language learning".

References

Bresnan, J. W. (1978). A realistic transformational grammar. In M. Halle, J. Bresnan & G. A. Miller (eds.), *Linguistic theory and psychological reality*, pp. 1–59. Cambridge, MA: MIT Press.

Carroll, S. (1989). Language acquisition studies and a feasible theory of grammar. *The Canadian Journal of Linguistics*, 34 (4), 399–418.

Carroll, S. (submitted). *Input and evidence: The raw material of second language acquisition*. MS: Universität Potsdam (1996).

Chomsky, N. (1965). *Aspects of the theory of grammar*. Cambridge, MA: MIT Press.

Green, D. W. (1986). Control, activation, and resource: A framework and a model for the control of speech in bilinguals. *Brain and Language*, 27, 210–23.

Pawley, A., & Hodgetts Syder, F. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (eds.), *Language and communication*, pp. 191–225. London: Longman.

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

Bilingualism: Language and Cognition 1 (1998) 25–26 © 1998 Cambridge University Press

Does the formulator know its LFG?

KEES DE BOT
Department of Applied Linguistics, Research Methodology and Business Communication Studies, University of Nijmegen, Erasmusplein 1, 6525 HT Nijmegen, The Netherlands.
E-mail: *c.debot@let.kun.nl*

A problem when reacting to a contribution as rich and as highly interesting as Pienemann's paper, is that it is much easier to react to what's not in it than to what is. An example is the relation between the theory proposed and the way linguistic knowledge is acquired: "[Processability theory] . . . is not designed to contribute anything to the question of the innate or learnt origin of linguistic knowledge or the inferential processes by which linguistic input is converted into linguistic knowledge." It is not easy to interpret the full consequences of the model if we are not to take into account the inferential processes by which the learner adds new information to his/her system. Maybe here the dividing lines are drawn too strictly: if the hypothesis space of the learner is restricted by the procedures that can be applied, one wonders at what level in the cognitive system these restrictions are located. Does the conceptualizer already know what the system is capable of doing, or is it uninformed (as follows from a strict interpretation of Levelt's proposal of an incremental system without feedforward)? In the latter case, this would mean that the learners/speakers would still try to express their communicative intentions and only later on in the process would things go wrong, for example, because the S-procedure cannot be carried out. In Pienemann's view this breakdown would lead to a return to a "direct mapping of conceptual structures on to surface form as long as there are lemmata that match the conceptually instigated searches of the lexicon." In terms of the model this means that an error message is sent to the conceptualizer with the intent that a different structure be started. How the conceptualizer in its ignorance of what's available on the next level knows how to ask for the right structure and procedures is a general problem for the Levelt model and not specific to Pienemann's use of it. My problem is that I fail to understand why the formulator would send this error message. Second language learners have all sorts of knowledge that they can use to solve problems that are not different in kind from the problems native speakers will have to deal with as well. If in a given situation the formulator is not able to generate the best solution, it will go for the second best. If for instance a V-procedure asks for three arguments (i.e. the learner's version of this verb has the conceptual arguments X,Y,Z and the grammatical functions SUBJ, DO, IO), and only two can be found in time, the system will not collapse and return to more primitive procedures: it will accept this less than perfect solution under the time pressure that is characteristic of language production. Vosse and Kempen (1991) present a very elegant model that may be capable of dealing with this issue. They tested a new cognitive architecture of human sentence processing (the unification space) against experimental data on dependency constructions in German and Dutch. The Unification Space as a metaphor is inspired by the metaphor of biochemical synthesis: "Segments as molecules are floating around in a test-tube and entering into chemical bonds with other molecules. The resulting larger structure may be insufficiently stable and fall apart again. After that the segments continue their search for suitable unification partners until a stable configuration – that is, the final parse tree – has been reached." Lemmas that are activated are entered in the Unification Space. The unification of two nodes can unify with a probability p(u). The probability will (among other things) depend on the grammatical "goodness of fit." The probability increases as the activation of the nodes and/or their grammatical goodness of fit decrease. "One consequence of this scheme is a bias in favor of semantically and syntactically well-formed syntactic trees encompassing recent nodes." High quality parse trees are "freezed." The unifying will have to lead to one frozen parse tree within a given time space. If there is not enough p(u) no configuration will emerge. This proposal is based on a native command of a language. It is very likely that most foreign languages learners will have to set the level of p(u) on a lower level than native speakers, or that when a problem is encountered and the optimal solution cannot be realized, the second-best parse is accepted. The unification space also seems to offer a processing based solution for the mechanism of feature matching that is central in Pienemann's theory, but that is only discussed in terms of LFG.

This brings me to my second point. Pienemann's intention is to develop a "processing-oriented approach to explaining language development." In order to achieve this, he combines theoretical notions from recent models of language production and notions from grammatical theories, in particular Lexical Functional Grammar. I am not sure that the balance between the contributions from the psycholinguistic models and the grammatical theories is satisfactory. In the language production models used, LFG is applied as the grammatical theory ("The architecture of LFG coincides with most of the key points . . . [related to] . . . language processing"). In the second part of Pienemann's paper one gets the feeling that the relation between LFG and processing has been inversed, however: the line of argumentation seems to be that because the developmental trends in second language learners can be explained using LFG notions, this is how processing takes place. I seriously doubt that that is sound reasoning, but maybe I misinterpreted Pienemann in this.

There are many more points of discussion, but that doesn't imply a lack of appreciation of Pienemann's con-

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

tribution. It will have a major impact on our thinking about the relationship between psycholinguistic knowledge, grammatical theory and SLA.

References

Vosse, T., & Kempen, G. (1991). A hybrid model of human sentence processing: Parsing right-branching, center-embedded and cross-serial dependencies. *Proceedings 2nd International Conference on Parsing Technologies.* Cancun, Mexico.

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

Bilingualism: Language and Cognition 1 (1998) 27–28 © 1998 Cambridge University Press

Semantic-informational and formal processing principles in Processability Theory

JAN HULSTIJN
Department of Applied Linguistics, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands.
E-mail: *hulstijn@let.vu.nl*

Let me begin my comments on Pienemann's keynote paper by expressing my admiration for the scholar who has developed Processability Theory (PT) over a period of some fifteen years with great determination and perseverance. What in earlier publications (e.g. Pienemann, 1985, 1987) appeared to me to be a rather disparate set of principles aiming to account for a limited set of empirical data (the well known sequence of five word orders of the ZISA study), has now evolved into a coherent theory which meets the demands of falsifiability, as PT's claims are formulated in sufficient detail to allow SLA researchers to put them to empirical test. PT comprises a number of principles of great generality, accounting, in principle, for the acquisition of any structure in any language, thereby exceeding the limits of the ZISA data of natural German L2 acquisition. As such, it is to be hoped that PT will have a healthy influence on the field of SLA research, as this field, in the last few years, has perhaps been dominated too much by the issue of whether L2 learners have access to Universal Grammar.

Now that PT has grown to full stature, many of the concerns I raised ten years ago (Hulstijn, 1987) have vanished. Yet I still have some reservations concerning the claimed psychological plausibility of PT, which I will specify in the remainder of this commentary.

(1) It is not clear to me what PT's claims are concerning the acquisition of subject-verb agreement (SV AGR) in German L2. Pienemann's paper at some points suggests that SV AGR can be processed (i.e. acquired) before verb separation (SEP) and inversion (INV), but at other points suggests that it can be processed and acquired after SEP, simultaneously with INV. The first claim runs counter to PT itself and the second is hard to accept on cognitive grounds. Saying that SV AGR can be acquired before SEP amounts to saying that feature unification at the clause level can be acquired before feature unification at the phrase level, which is at variance with one of the central claims of PT. On the other hand, the claim that SV AGR can only be acquired after SEP runs into other difficulties. From a cognitive perspective, in contrast to a linguistic perspective, it might be argued that the difference between "das Kind spielt" (the child plays) and "die Kinder spielen" (the children play) is a semantically and communicatively more important distinction than the semantically and communicatively void difference between VPs such as "muß machen die pause" (must make the break) and "muß die pause machen" (must the break make). On cognitive grounds, one might therefore expect learners to implement correctly SV AGR before implementing SEP. I wonder whether Pienemann could inform us what his German L2 data (both the cross-sectional ZISA data and the longitudinal data collected among Australian learners of German) reveal in this respect. Did learners produce SV AGR in SVO sentences before they produced SEP? My expectation is they did. Similarly, why is it that learners produced INV in WH questions before they did so in declarative sentences beginning with an adverbial phrase? Isn't this because INV in the former case is semantically motivated, whereas INV in the latter case is an exclusively syntactic procedure? Shouldn't the (triggering?) role of semantics then be more fully integrated into PT?

(2) All processing prerequisites in PT except one are formal in nature. These are the lemma, category, phrase, sentence, and the subordinate clause procedures. PT invokes only one non-formal principle, namely perceptual salience. This is done to account for the occurrence of adverbials at sentence initial or final position (ADV). Being the only non-formal principle, perceptual salience in PT has the appearance of an ad hoc principle. It seems to me, however, that a theory which aims at providing a full explanation of SLA should be based on both formal (morphosyntactic) and "informational" principles, which are to be integrated in a developmental pattern. Such an approach to language development calls for a competition model. In the earlier stages of language development, informational principles will rule, as it were, with absolute power (as has been demonstrated by researchers involved in the ESF project; see, for instance, Klein and Perdue, 1992). In later stages, various formal principles are acquired (perhaps in the order as outlined in PT), limiting the possible forms utterances can take. However, even at the final stage of acquisition, utterances are never exclusively determined by formal principles (e.g. in the linear ordering of topicalized and focused information). This calls for a theory specifying the time course of the competition between semantic-informational principles on the one hand and formal processing principles on the other.

In conclusion, PT now deserves to be called a true SLA theory. Its author is to be complimented on it, and the field of SLA research should take note, as PT contains many falsifiable claims and hence has the potential to affect SLA research in a fruitful way. In my estimation, however, PT has not entirely succeeded yet in explaining how language learners cope with informational and linguistic demands at various stages of language development. The role of non-formal principles, such as perceptual salience, deserves to be fully integrated into PT, allowing semantic-informational and formal principles to compete with each other in different ways at different stages of development. I hope Pienemann will strive to develop further PT along this line.

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

References

Hulstijn, J. (1987). Onset or development of grammatical features? Comments on Pienemann's approach to acquisition orders. Paper presented at a workshop on SLA, held at La Trobe University (Melbourne, Australia, 1987).

Klein, W., & Perdue, C. (1992). *Utterance structure*. Amsterdam: John Benjamins.

Pienemann, M. (1985). Learnability and syllabus construction. In K. Hyltenstam & M. Pienemann (eds.), *Modelling and assessing second language acquisition*, pp. 23–76. Clevedon, UK: Multilingual Matters.

Pienemann, M. (1987). Determining the influence of instruction on L2 speech processing. In G. Kasper (ed.), *Classroom Research. AILA Review*, 5, 40–72.

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

Bilingualism: Language and Cognition 1 (1998) 29–30 © 1998 Cambridge University Press

Comparing and explaining the trajectories of first and second language acquisition: in search of the right mix of psychological and linguistic factors

GERARD KEMPEN
Experimental and Theoretical Psychology Unit, Leiden University, PO Box 9555, 2300 RB Leiden, The Netherlands.
E-mail: *kempen@rulfsw.leidenuniv.nl*

When you compare the behavior of two different age groups which are trying to master the same sensori-motor or cognitive skill, you are likely to discover varying learning routes: different stages, different intervals between stages, or even different orderings of stages. Such heterogeneous learning trajectories may be caused by at least six different types of factors:

(1) Initial state: the kinds and levels of skills the learners have available at the onset of the learning episode.
(2) Learning mechanisms: rule-based, inductive, connectionist, parameter setting, and so on.
(3) Input and feedback characteristics: learning stimuli, information about success and failure.
(4) Information processing mechanisms: capacity limitations, attentional biases, response preferences.
(5) Energetic variables: motivation, emotional reactions.
(6) Final state: the fine-structure of kinds and levels of subskills at the end of the learning episode.

This applies to language acquisition as well. First and second language learners probably differ on all six factors. Nevertheless, the debate between advocates and opponents of the Fundamental Difference Hypothesis concerning L1 and L2 acquisition have looked almost exclusively at the first two factors. Those who believe that L1 learners have access to Universal Grammar whereas L2 learners rely on language processing strategies, postulate different learning mechanisms (UG parameter setting in L1, more general inductive strategies in L2 learning). Pienemann opposes this view and, based on his Processability Theory, argues that L1 and L2 learners start out from different initial states: they come to the grammar learning task with different structural hypotheses (SOV versus SVO as basic word order of German).

However, acquisition routes may diverge in response to the other types of causal factors as well. Language input, which may vary widely between language learning situations, is a case in point. For example, M. Kempen, Gillis and Wijnen (in press) have observed that, in Dutch, child-directed speech contains an abundance of sentence-final non-finite verbs with very transparent meanings (referring to concrete actions). They suggest that this input feature renders such verbs much more salient than finite verbs in sentence-initial positions. This saliency, they argue, is one of the factors contributing to the high frequency of non-finite verbs at the end of the children's own utterances. Although I have no pertinent data at hand, it seems likely that similar characteristics are less prominent in the language input received by L2 learners of Dutch. On the assumption that these contingencies hold for German as well, the prediction follows that SOV as an initial structural hypothesis is more likely in L1 than in L2 learners of German.

This contrast may be sharpened by additional differences between L1 and L2 learners as regards their information processing capabilities (cf. the fourth factor in the above list). In comparison with young children, L2 learners have available a more diversified "lexicon" of concepts, including concepts not only for concrete actions but also for abstract states, intentions, temporal relationships, etc. This helps them to recognize, in the language spoken in their environment, verbs designating such abstract concepts – among them the modal and copula verbs. These verbs often occupy early positions in a sentence. Wijnen (in press) and M. Kempen et al. indeed observed that the first finite verbs produced by Dutch children in sentence-initial positions are not the inflected forms of the non-finite verbs they have acquired earlier (i.e. those designating concrete actions), but originate from a novel vocabulary of "non-eventive" (abstract, state) verbs. Since L2 learners have already acquired the abstract concepts underlying the latter verbs at the onset of the learning episode, they are expected to discover the verb-second rule of German at an earlier stage than L1 learners.

Arguments and observations like the above do not provide a complete account of the differences between the German L1 and L2 acquisition trajectories discussed by Pienemann. What they do bring out is the role that non-syntactic factors play in the process of acquiring the syntactic rules of a language. This role should be acknowledged in accounts that appeal to syntactic factors, whether in terms of Universal Grammar or Processability Theory.

Pienemann is right in pointing out that the explanations in terms of (limited) access to Universal Grammar that have been proposed for the two acquisition trajectories are unsatisfactory (see the third section). However, his own account based on Processability Theory is not unproblematic either. In the first section (subsection entitled "Principles of processability"), Pienemann assumes that word order of the phrases that have been attached to an S-node, is arranged by a specialized processing module called S-procedure. One of the word order rules (introduced in the second section, "second language development") is R1. It expresses the structural hypothesis (SVO) that L2 learners

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

are supposed to entertain initially, that is, at the first acquisition stage. This implies that the S-procedure must be available at that early point in time. However, at the end of the second section, Pienemann writes "that the S-procedure is hypothesised to become available at level 5 of the processability hierarchy", in fact, close to the end of the learning episode. This assumption, which also figures in Table 1, is needed in order to explain the relatively late appearance of Subject-Verb inversion in L2 utterances (and, in the third section, for the late emergence of finite verbs and their verb-second placement). I cannot escape the conclusion that, in the present version, Processability Theory is internally inconsistent.

If asked my personal opinion of the explanation of similarities and differences between L1 and L2 acquisition trajectories, I would suggest that psychological factors such as those revealed by the studies cited above deserve serious consideration.

References

Kempen, M., Gillis, S., & Wijnen, F. (in press). Dutch children and their mothers' infinitives. *Proceedings of GALA 1997*.

Wijnen, F. (in press). Functionele categorieën in Nederlandse kindertaal. *Nederlandse Taalkunde*.

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

Bilingualism: Language and Cognition 1 (1998) 31 © 1998 Cambridge University Press

How to slice the cake?

PIETER MUYSKEN
Department of Linguistics, Universiteit van Amsterdam, Spuistraat 210, 1012 VT Amsterdam, The Netherlands.
E-mail: *pieter.muysken@hum.uva.nl*

Manfred Pienemann offers an interesting, fair-minded, and systematic account of the role of processability constraints in explaining language development. Thus he opens up a wide field of research in the area of the interaction of grammatical knowledge and processing constraints, and Pienemann should be applauded for this kick-off. In my view, this field should be explored in a non-partisan manner. Rather than pushing either grammar or processing at the expense of the other, we should be looking at the division of labour between different mental capacities, and at potential sites of overlap, conflict, and convergence. L2 development is crucial in this research, because it may show how much can be accomplished with non-grammatical processing. I will limit myself here to a number of disjointed remarks on some directions for research.

The notion of generative entrenchment is very useful, I think, to explain the straight course of first language development. However, for adult second language learning, I am less certain. First, early Turkish learners of Dutch and German show evidence of a non-finite OV stage, which is then replaced by a fairly rigid SVO-stage. Second, learners may vary in their ultimate attainment of the SOV+V2 pattern. Third, I would like to see whether the amount of variation between competing order patterns in L2 development is not greater than the "interlanguage grammar" summaries would suggest. The notion of "interlanguage grammar" made L2 research academically respectable in the linguistic community in the seventies (in the same way as "UG-based L2" did in the eighties), but the notion should be held to the light critically, I think.

Zwart (1993) and others have been arguing that the traditional generative accounts of Dutch and German as SOV+V2 systems were flawed from the outset. (Thus generative entrenchment can also explain the emergence of "received wisdom".) In fact, they claim that Dutch and German are basically VO, an analysis in line with Kayne's (1994) proposal that all languages are underlyingly VO. It would be worthwhile to review once again both the L1 and L2 data in terms of the specific VO features pointed at in this literature. We may be in for some surprises.

Pienemann does not enter into a discussion of the degree to which processing is language-specific. It is clear that the proposed processability hierarchy is universal. However, this may not extend to specific strategies within this hierarchy. Then the question becomes to what extent processing strategies developed for the first language operate in interpreting second language input. Again, this is an important area for future research.

Related to this is the unresolved question of age. Since Pienemann does not allude to the specific role of the L1, his account is roughly compatible with one in terms of early and late development. Is there a qualitative jump from Meisel's bilingual children, behaving like L1 learners in this respect, to Pienemann's own young L2 learners, behaving like adults? Where is the cut-off point between the two groups? Is it age, or rather sequential (and not simultaneous) acquisition? Should both groups, Meisel's and Pienemann's, not be re-examined in the light of the data from their counterparts?

In summary, the L1/L2 German word order facts, rather than being a dead-end, continue to pose a number of interesting and fundamental questions.

References

Kayne, Richard S. (1994). *The anti-symmetry of syntax.* Cambridge, MA: MIT Press.

Zwart, C. Jan Wouter (1993). *Dutch syntax. A minimalist approach.* Dissertation, Groningen University.

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

 Bilingualism: Language and Cognition 1 (1998) 32–33 © 1998 Cambridge University Press

The need for converging evidence

JACQUELYN SCHACHTER
Department of Linguistics, University of Oregon, Eugene, OR 97403, USA. ***E-mail:*** *schachte@oregon.uoregon.edu*

Pienemann's Processability Theory is an interesting and well thought out attempt to account for the developmental problem in L1 and L2. If correct, it would relieve the genetically specified knowledge base of what he views as its burdensome task of having to specify a great deal of genetically encoded linguistic knowledge and transfer at least part of that task to the processing mechanisms available for the acquisition of procedural skills needed for language processing. He posits a set of processing resources forming an implicational hierarchy such that the resource of a certain level is a prerequisite for the functioning of the next higher level. He posits further that the processing devices will be acquired in their predetermined sequence of activation in the production process. A consequence of this is that acquirers of both the first and the second language will follow the same sequence of acquisition of the processing hierarchy. The differences in the courses of development can be attributed, he says, to the knowledge base with which the learner comes to the acquisition task. In the case of L1, he says, many would see UG as the knowledge base; in the case of L2, only the surface syntactic features of the target language are available (and, presumably, knowledge of the L1).

With an interesting proposal such as this, it would be important to look for converging evidence from other fields of inquiry that would serve to substantiate some of Pienemann's claims. One field of interest is that of the neuroimaging of the activation of the various areas of the brain while processing language. Such imaging (using PET scans, ERPs, and fMRIs) has begun to amass valuable information about what parts of the brain are active when the subject is carrying out various different kinds of linguistic tasks.

It would seem, for example, that if the language processor develops in the same way for both child L1 and adult L2 acquisition, then there should be some spatially located area of the brain that processes the same tasks in the same way. That is, we wouldn't expect to see evidence that one area of the brain is activated for a particular syntactic task in the case of the monolingual adult L1 speaker and yet another area is activated for the same linguistic task in the case of the adult L2 speaker of the same language (acquired as an adult). Yet that is what appears to be the case in the imaging studies so far completed.

Turning to the ERP literature we find some that is of special interest to the language acquisition theorist modeling language acquisition. As Weber-Fox and Neville (1996) note, "recent ERP evidence suggests that . . . the ERP responses for at least some kinds of syntactic processing [in monolinguals] may be generated within anterior regions of the left hemisphere (Neville, Mills and Lawson 1992; Kutas and Kluender 1993, etc.)" (p. 4), whereas the lexicon is processed in the temporal-parietal/medial temporal region (cf. Ullman at al., 1997). That is, when adult monolingual native speakers of a language provide data on processing grammatical and ungrammatical sentence pairs, the results generally indicate that syntactic processing takes place in Broca's area. But when the same adult monolinguals process semantically nonanomalous and anomalous sentence pairs, the results indicate that semantic processing is usually distributed across both hemispheres and in posterior regions of the brain. Other studies indicate that the syntactic processing system is more sensitive to critical period effects than is the semantic processing system (Neville, 1990; Neville, Coffey, Lawson, Fischer, Emmorey & Bellugi, 1997). That is, although semantic processing appears to take place in the same anatomical area of monolingual and late-learning speakers of a particular language, syntactic processing by the same contrasting groups appears to take place in different areas. Weber-Fox and Neville (1996) have shown, for example, that cerebral organization in late-learned adult speakers of English is associated with reduced specialization in left hemisphere language processing systems and may include increased right hemisphere involvement.

This kind of evidence, new in the field of second language acquisition research, is nevertheless highly relevant. And it must be taken into account in model building. If syntactic processing is carried out in different regions of the brains of monolingual and late-learning bilingual speakers of the same language, it must be that they are using different mechanisms. For these reasons, I remain doubtful that Pienemann is on the right track.

References

Kutas, M., & Kluender, R. (1993). What is who violating? A reconsideration of linguistic violations in light of event-related brain potentials. In H. J. Heinze, T. F. Munte & G. R. Mangun (eds.), *Cognitive electrophysiology: Basic and clinical applications*, pp. 183–210. Boston: Birkauser.

Neville, H. (1990). Intermodal competition and compensation in development: Evidence from studies of the visual system in congenitally deaf adults. In A. Diamond (ed.), *The development and neural bases of higher cognitive function*, pp. 71–91. New York: New York Academy of Sciences Press.

Neville, H. J., Coffey, S. A., Lawson, D. S., Fischer, A., Emmorey, K., & Bellugi, U. (1997). Neural systems mediating American sign language: Effects of sensory experience and age of acquisition. *Brain and Language*, 57, 285–308.

Neville, H., Mills, D. L. & Lawson, D. S. (1992). Fractionating language: Different neural subsystems with different sensitive periods. *Cerebral Cortex*, 2, 244–58.

Swinney, D., Zurif, E., Prather, P., & Love, T. (1996). Neurological distributions of processing resources underlying language

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

comprehension. *Journal of Cognitive Neuroscience*, 8 (2), 174–84.

Ullman, M., Corkin, S., Coppola, M., Hickok, G., Growdon, J. H., Koroshetz, W. J., & Pinker, S. (1997). A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *Journal of Cognitive Neuroscience*, 9 (2), 289–99.

Weber-Fox, C. M., & Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience*, 8, 231–56.

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

 Bilingualism: Language and Cognition 1 (1998) 34–35 © 1998 Cambridge University Press

On the "wrong-headedness" of generative entrenchment*

BONNIE D. SCHWARTZ
Department of Linguistics and English Language, University of Durham, Elvet Riverside II, New Elvet, Durham DH1 3JT, UK.
E-mail: *B.D.Schwartz@durham.ac.uk*

The main thesis Pienemann puts forward is that while L1 acquisition and L2 acquisition of a particular language (e.g. German) exhibit non-parallel paths of development, such a difference does not entail that the processing resources claimed to be responsible for transitioning from one stage to the next are any different in the two situations. The cause of the difference in routes, he suggests, is the difference in departure points. In the German case at hand, L1 acquirers begin with an SOV grammar, while L2 acquisition, according to Pienemann, starts off SVO ("canonical order"). "Generative entrenchment" of SVO is what makes the ensuing development distinct from that of L1 German: "once a decision has been made and a new structure has been added, it is very costly, if not impossible, for the developmental process to move to a different developmental path." I think there's something very right about this general picture – but also that part of it is "wrong-headed."

Pienemann attempts to argue against the logical counterproposal to entrenched SVO, namely, that "L2 learners fundamentally restructure the developing grammar and switch to the assumption that German is an SOV language." This restructuring idea originated in du Plessis, Solin, Travis and White (1987) in regard to stage 3 (S $V_{[+fin]}$ O $V_{[-fin]}$) of the ZISA L2 sequence. However, Pienemann wonders "on the basis of which . . . evidence the learner would arrive at such a conclusion, since the SVO structures [sic] they produce form a subset of the input data." This is perhaps a clue that restructuring to SOV is not so outlandish, because the *same* holds of children acquiring German as L1 – "the SVO [orders German children] produce form a subset of the input data" – and we know German children *do* posit an SOV grammar. But in order to argue for L2 restructuring to SOV, such a clue is actually unnecessary; all we really need to do is examine whether "[n]either . . . PART [for stage 3] nor INV [for stage 4, XP $V_{[+fin]}$ S . . .] is compatible with the SOV assumption." So, reconsider the stage 3 c-structure rules:

(1) Stage 3, L2 German: S $V_{[+fin]}$ O $V_{[-fin]}$
(R3): $S \rightarrow NP_{subj}\ VP$
$VP \rightarrow V \left\{ \begin{matrix} (NP_{obj1})\ (NP_{obj2}) \\ \text{V-COMP} \end{matrix} \right\}$
$\text{V-COMP} \rightarrow (NP_{obj1})\ (NP_{obj2})\ V$

For Pienemann, "(R1) [$S \rightarrow NP_{subj}\ V\ (NP_{obj1})\ (NP_{obj2})$] *develops* into (R3)" (M. Pienemann p.c.). That is, what was originally a purely SVO system now includes a c-structure rule that *generates* OV: V-COMP. The introduction of V-COMP in the development of (R1) into (R3) thus *is* restructuring from VO to OV.

* Thanks to Rex Sprouse, my partner in crime.

Yet there's a difference between (R3) and the restructuring posited by duPlessis et al.: (R3) is a mixed system, i.e. it amounts to saying that every verb takes its complement to the right ($VP \rightarrow V\ (NP_{obj1})\ (NP_{obj2})$) and to the left ($\text{V-COMP} \rightarrow (NP_{obj1})\ (NP_{obj2})\ V$). So in a sentence with a simplex verb, as in (2a), "lesen" takes its complement to the right, but as the complement to an auxiliary, "lesen" takes its complement to the left ((2b)):

(2) (a) Ich lese EIN BUCH
I read a book
(b) Ich kann EIN BUCH lesen
I can a book read

Similarly for auxiliaries themselves – since they, too, serve as complements to other auxiliaries, as in (3) – and even for separable prefix verbs, as in (4), where here Pienemann would have to say that the prefix "an" is the "complement" to the verb (rufen):

(3) (a) Ich muß LESEN KÖNNEN
I must to-read be-able
"I have to be able to read"
(b) Ich habe LESEN KÖNNEN müssen
I have to-read be-able must
"I had to be able to read"
(4) (a) Ich rufe AN
I call up
(b) Ich muß ANrufen
I must up-call

Whether all this is feasible is perhaps a matter of theoretical taste. Whether it's empirically adequate is another story, as we shall now see.

Recall that (R3) is never replaced by any other c-structure rule. At stage 4 a new c-structure is added for INV (XP $V_{[+fin]}$ S . . .), and at stage 5 this rule is itself modified by a "constraining equation," whose intended purpose is restricting V2 to matrix clauses. Ruling out V2 *in* embedded clauses, however, has no bearing on the position of the verb that (R3) will generate *for* them. In other words, (R3) incorrectly predicts that at stage 5, L2ers will (still) produce, embedded under a root clause, only sentences realizing exactly those orders discussed in (2)–(4), for instance, (5) – cf. (2):

(5) (a) Hans sagt, daß ich lese EIN BUCH
Hans says that I read a book
(b) Hans sagt, daß ich kann EIN BUCH LESEN
Hans says that I can a book read

But what (R3) cannot generate is "Hans sagt, daß ich ein Buch LESE" and "Hans sagt, daß ich ein Buch lesen KANN" – precisely what stage 5 is about: embedded finite verbs in clause-final position.

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

Summarizing: contrary to generative entrenchment, which bars (or at least strongly disfavors) restructuring, Pienemann's analysis itself embodies just such a property, at least partially – i.e. restructuring from VO to both VO and OV. Yet this partial restructuring turns out to be empirically inadequate for the embedded context of stage 5. What's needed is full restructuring of the L1 VP, from VO to OV. This in conjunction with one other change captures the verb-final facts of stage 5 (on the "change" of stage 5, see du Plessis et al., 1987 for one proposal and Schwartz & Tomaselli, 1990 for another).

While generative entrenchment does not seem viable, I nevertheless think it would be misguided to deny Pienemann's insight: that the developmental course of language is fundamentally steered by the initial hypothesis assumed on the part of the acquirer, and furthermore that because the L1 and L2 initial hypotheses are distinct, distinct developmental paths ensue. This leads, of course, to the question Pienemann himself left us with: what is the source of the L2er's initial hypothesis? Or to be more specific, what is the source of the initial SVO hypothesis of the ZISA informants? Or more generally, would all L2ers approach German with the assumption of SVO? These questions have been much debated in the literature. On my reading, the evidence overwhelmingly suggests that the source of initial L2 hypotheses is the L1 grammar, and so the reason the ZISA subjects' initial hypothesis about German is SVO is because Italian, Portuguese and Spanish are SVO, while the reason, for example, the lowest-level Koreans and Turks in the Vainikka and Young-Scholten (1994) study rely so heavily on SOV is because their mother tongue is SOV.

Now, if the sole property of grammar determinant to course of development were basic word order, one would expect developmental paths to converge when the basic word order – either by initial hypothesis or by subsequent restructuring – coincides with that of the target language. But the evidence does not support this: for example, the longitudinal (L1 Turkish) L2 German data examined by Schwartz and Sprouse (1994) clearly exhibit OV at all times; yet, the main-clause XP $V_{[+fin]}$ S . . . order develops from nonexistent, to with pronominal subjects only, to, finally, with both pronominal and nonpronominal subjects (see Schwartz & Sprouse for details and an attempt to address the "transition problem") – quite unlike L1 development of German. So it cannot be just the basic word order of the L1 that constitutes the initial L2 assumption; more is needed. Indeed, the hypothesis Schwartz and Sprouse (e.g. 1994, 1996) have offered and defended is that the L2 initial state comprises much, much more of the L1 grammar: the L1 in its entirety. This predicts that in the L2 acquisition of a given language, the extent to which the L1s differ is the extent to which the L2 developmental paths will likewise differ. The L2 initial state will see its effects throughout the L2 acquisition process, because

> the cause of *any* (developmental) differences in the L2 acquisition of a particular language that covary with native language *must* be present from the beginning (namely, must be the result of the L2 initial state), precisely because the cause of such covariation cannot be in the input, since this remains constant. (Schwartz & Sprouse, 1996: 67)

To conclude, although the "points of departure" for Pienemann and for Schwartz and Sprouse are distinct (transition theory and property theory, respectively), that their approaches converge on the propagation effects of the initial L2 hypothesis is, I believe, a very promising sign for the field.

References

Du Plessis, J., Solin, D., Travis, L., & White, L. (1987). UG or not UG, that is the question: A reply to Clahsen and Muysken. *Second Language Research*, 3, 56–75.

Schwartz, B. D. & Sprouse, R. A. (1994). Word order and nominative case in nonnative language acquisition: A longitudinal study of (L1 Turkish) German Interlanguage. In T. Hoekstra & B. D. Schwartz (eds.), *Language acquisition studies in Generative Grammar*, pp. 317–68. Amsterdam: John Benjamins.

(1996). L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research*, 12, 40–72.

Schwartz, B. D., & Tomaselli, A. (1990). Some implications from an analysis of German word order. In W. Abraham, W. Kosmeijer & E. Reuland (eds.), *Issues in Germanic syntax*, pp. 251–74. New York: Mouton de Gruyter.

Vainikka, A., & Young-Scholten, M. (1994). Direct access to X'-theory: Evidence from Korean and Turkish adults learning German. In T. Hoekstra & B. D. Schwartz (eds.), *Language acquisition studies in Generative Grammar*, pp. 265–316. Amsterdam: John Benjamins.

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

 Bilingualism: Language and Cognition 1 (1998) 36–38 © 1998 Cambridge University Press

AUTHOR'S RESPONSE

A focus on processing

MANFRED PIENEMANN
Linguistics Program, Department of Modern European Languages, Australian National University, Canberra, ACT 0200, Australia.
E-mail: *Manfred.Pienemann@anu.edu.au*

Given that my paper touched upon many aspects of L1 and L2 acquisition, it is not surprising that the commentaries have raised a diverse range of issues. Nevertheless, one should not lose sight of the fact that the paper has a tightly defined focus, namely to delineate the influence of language processing factors and developmental dynamics on developmental trajectories in L1 and L2 acquisition, and this is done in a very explicit manner. As Carroll rightly puts it in her commentary, this "[e]xplicitness allows us to formulate questions which could otherwise not be asked."

One has to accept, however, that the virtue of explicitness comes at a price. One cannot achieve explicitness and completeness at the same time, at least not in a new field such as SLA. This put me in the same dilemma as several of the commentators. In designing a novel approach to language acquisition should one aim for including all factors that may be relevant, and thereby compromise on explicitness, or should one aim for explicitness in one theory component and thereby guarantee its testability while foregoing completeness? I chose the second avenue.

This nevertheless allows me to agree in principle with many of the comments which aim at completeness. For instance, I agree with Kempen that his six key factors are likely to contribute to the explanation of L1–L2 developmental differences. However, such a global commitment is quite a different matter from working out detailed theory components that make the mechanics of the influence of each of these factors and their interaction explicit. Kempen's "energetic" variable is a point in case.

I also agree with Carroll who suggests, amongst other things, that a plausible theory of SLA needs to include a component which accounts for the acquisition of linguistic skills. I further agree with Hulstijn that a theory component needs to be included which deals with the competition between formal and semantic-informational principles.

As I said above, I chose a reductionist and explicit approach which focuses on the impact of processing factors on developmental trajectories. The additional theory components suggested by the commentators will no doubt keep a generation of researchers busy if those also aim for explicitness.

Pieter Muysken appears to be quite aware of this dilemma when he adds a few more items to the list of explananda: typological plausibility, transfer of processing mechanisms and the critical period. Some of these issues are in fact dealt with in Pienemann (in press).

In summary, I dare to draw a positive conclusion from the observed incompleteness and explicitness of Processability Theory. In Muysken's words, "[It] opens up a wide field of research."

Let me turn to specific points in individual commentaries. De Bot points out a limitation of Levelt's model and proposes Vosse and Kempen's (1991) model that overcomes this problem in principle. I agree that this is a very promising direction, and one that I am interested in pursuing in the future. However, the process of adapting this model to a wide range of typologically diverse developmental phenomena is not a trivial one and will require a very intense research effort.

Hulstijn raises concerns about the developmental relationship of morphological structures (e.g. subject-verb agreement) and word order (INV, SEP, etc). A similar point was raised by one reviewer. Let me repeat the corresponding paragraph from the paper. "One reviewer commented that one might conclude from this that Processability Theory entails the prediction that the above structures [INV and subject-verb agreement] are acquired simultaneously. However, it would be incorrect to derive such a prediction from the theory, since it does not predict that whatever can be processed will indeed be acquired. Instead, the theory predicts that what cannot be processed will not be acquired. In other words, processability acts as a constraint on development and therefore does not entail the prediction that the above structures are acquired simultaneously." I mentioned in the paper that we have indeed found that the developmental schedule of syntax and morphology does not interlock in child or adult SLA. The other side of the coin is the notion of "developmental gaps." In the context of SLA, I have shown experimentally (Pienemann, 1988) that structures which are processable but absent from the interlanguage can be added to it through formal intervention.

Kempen raises briefly his concern that the S-procedure is present in rule 1 of my framework while it appears only at level 5 of the processability hierarchy. I have to admit that my presentation may have been misleading here. I use LFG notations primarily as a convenient tool to describe the unification of lexical features and word order constellations, not to map all procedures. I assume that the S-procedure is initially unable to retain information on lexical features, and when it appears at level 5 it does have this capacity.

Carroll wonders why nothing hinges on the adherence to LFG. The answer is that Processability Theory is not a property theory, but a transition theory. It focuses on the question of which grammatical forms are processable at different stages of development, not on how the properties of the related grammars are arrived at. Also note that, contrary to Carroll's assumption, notions of linguistic memory and language processing are not (in Processability

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

Theory) inferred from LFG or any other linguistic theory, but from research on language processing, and LFG is used as a convenient short-hand notation of the crucial lexical unification processes involved in language processing. This notation has a high degree of psychological plausibility (cf. Pienemann, in press). An added advantage of using LFG in this context is that it is sufficiently general to be applicable to a wide range of linguistic phenomena across typologically different languages and it does allow for a future extension of Processability Theory into issues relating to a property theory. At that stage, the added theory component would naturally be very closely linked with an adherence to the chosen grammatical theory.

While Carroll found not enough adherence to LFG, Bialystok is critical about too much of it. The latter commentator wonders how "process" is defined in my paper. In her view, my paper merely ". . . acknowledg[es] a theoretical lineage to Levelt (1989) and others . . . ." With all respect, that is incorrect. The paper spells out over about one third of its length key psychological factors in *language* processing and how these support the basic architecure of the model proposed by Levelt (1989). It further describes in detail the generation of linguistic structures as an incremental process within Kempen and Hoenkamp's (1987) approach, on which components of Levelt's model are built. From this it is clear that the paper is about online incremental language processing in the productive mode. And since Processability Theory deals with the processability of linguistic structures, it naturally requires a psychologically plausible notational system which can generate precisely those linguistic structures which can or cannot be proessed by the learner. No matter what the architecture of the language generation process is, it would always need to be capable of generating the same structural output. Bialystok is concerned how ". . . one distinguish[es] between process and its attendant structure . . . ." The answer is that these are kept quite separate in the chosen framework for language generation. To quote the first secion of my paper, "In this perspective the language processor is seen with Kaplan and Bresnan (1982) as the computational routines that operate on (but are separate from) the native speaker's linguistic knowledge." This can be illustrated with one of the key processes which Processability Theory builds on, namely feature unification. Kempen and Hoenkamp's (1987) incremental procedural grammar spells out the procedures required to execute this process. Feature unification is not a grammatical structure itself, but it results in grammatical structures. In addition, it requires the presence of certain bits of information in the mental lexicon which in a theory of grammar translate into linguistic knowledge.

A brief comment on Bialystok's "problem of evidence." Historically, Bialystok is correct. The ZISA data were there before the explanation. However, the original explanatory framework was independently motivated. In the intervening twenty years, the explantory framework has been transformed into Processability Theory which has been tested on typologically different languages. In Pienemann (in press) I have shown that the processability hierarchy makes accurate predictions for the development of Swedish, Japanese and English as L2s, and work is under way on Mandarin, Spanish, Italian and Indonesian. The processability hierarchy has also been shown to make correct predictions in reaction time experients (Pienemann, in press). All this confirms the independence of the hierarchy from the first set of data.

Schwartz contrasts her own position with that put forward in my paper where I argue that the initial structural hypotheses in L1 and L2 get generatively entrenched and that grammar is not restructured in this developmental process. Schwartz reasons that the development of R1 into R3 *is* restructuring. In addition she argues that there is an error in the proposed rule system and that it really requires ". . . a full restructuring of the L1 VP from VO to OV."

Let me address the second point first. In the full exposition of German word order development within Processability Theory in Pienemann (in press), I assume the following c-structure rule which is added to the system at the stage V-Final:

$$S \rightarrow (COMP)_{ROOT=-}\, NP_{subj}(NP_{obj1})(NP_{obj2})(ADJ)(V)_{INF=-}\, V_{INF=+}$$

The discussion of the equation INF=+ in the second section of my paper implies such a rule, and this represents the empirical facts quite accurately.

This leads us to Schwartz's first point. She maintains that the development of R1 into R3 constitutes a case of restructuring, and one would assume that she would also count the addition of the above c-structure rule as a case of restructuring. However, my point is that R3 has evolved from R1 and contains R1, and when the above rule is introduced, R3 is maintained. In other words, new rules are added to the system. Nothing gets thrown away. This is the point behind generative entrenchment. Patterns are preserved and added on to. This structural conservation has the enormous computational advantages which I pointed out in the paper.

This is quite different from Schwartz's own view of development which assumes a fundamental change of the grammatical system (Schwartz and Tomaselli, 1990). The VP changes from VO to OV, and all structures are generated on this basis. This change amounts to throwing the old VP out when the new one comes in. Such a process would be totally discontinuous and contradicts the priniple of pattern conservation. In the paper I argued that such a fundamental restructuring would be computationally very costly.

Schachter raises the interesting point of converging evidence from neuroscience. In principle, I agree with her that such evidence has the potential to add to the plausibility of the psychological line of argument (cf. also Bechtel, 1988). Schachter quotes studies which do in fact support quite a number of assumptions concerning the architecture of the language processor which are made in conjunction with Processability Theory. An example is the independent and separate status of the lexicon and grammatical processing.

Her key point, though, is the study by Weber-Fox and Neville (1996) which demonstrates that adult L2 learners

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press

show ". . . reduced specialization in left hemisphere language processing systems and may include increased right hemisphere involvement" (Schachter). From this Schachter concludes that if L1 and L2 learners of the same language use different regions of the brain for syntactic processing then there must be different mechanisms for L1 and L2.

The question of lateralisation in bilinguals is one that has received a great amount of attention since the work of Albert and Obler (1978). A wide range of research techniques have been used in this context: electroencephalograms, tachistoscopic tests, dichotic listening tests, eye movement studies, time-sharing studies, aphasia studies (cf. Mägiste, 1992; Hoosain, 1992) and more recently the ERP (event-related potentials) studies mentioned by Schachter. The overall picture which emerges from these studies is contradictory. Already Albert and Obler (1978) suggested that linguistic functions in bilinguals are less lateralised and that the right hemisphere is more involved, particularly in the L2. However, this view has been challenged on empirical grounds, among others, by Paradis (1990). Reflecting the contradictory nature of the findings, the title of his paper is "Language lateralization in bilinguals: Enough already!"

The one study by Weber-Fox & Neville (1996) which Schachter quotes does not arbitrate on this issue. I am also somewhat concerned about the way in which linguistic and neurological structures appear to be seen in a one-to-one relationship. Schachter refers to "left hemisphere language processing systems" as if the neurological functioning of the language processor had been spelt out. However, one has to bear in mind that the relationship between neurological and linguistic structures is at this stage still a relatively indirect one (cf. Kolb and Whishaw, 1995).

Finally, allow me to refer back to the aforementioned statement in Muysken's commentary (". . . he opens up a wide field of research"). It is indeed my hope that this contribution may open up a wider field of research than would be accessible with an exclusive focus on Universal Grammar. Much work remains to be done.

References

Albert, M. L., & Obler, L. K. (1978). *The bilingual brain: Neuropsychological and neurolinguistic aspects of bilingualism*. New York: Academic Press.

Bechtel, W. (1988). *Philosophy of Science: An overview for cogitive science*. Hillsdale, N.J.: Lawrence Earlbaum.

Hoosain, R. (1992). Differential cerebral laterisation of Chinese-English bilingual functions? In R. J. Harris (ed.), *Cognitive processing in bilinguals*, pp. 561–71. Amsterdam: North Holland Elsevier.

Kaplan, R., & Bresnan J. (1982). Lexical-Functional Grammar: A formal system for grammatical representation. In J. Bresnan (ed.), *The mental representation of grammatical relations*, pp. 173–281. Cambridge, MA: MIT Press.

Kempen, G., & Hoenkamp, E. (1987). An incremental procedural gramar for sentence formulation. *Cognitive Science*, 11, 201–58.

Kolb, B., & Whishaw, I. Q. (1995). *Fundamentals of human neurocience*, 4th edn. New York: W. H. Freeman & Co.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Mägiste, E. (1992). Learning to the right: Hemispheric involvement in bilinguals. In R. J. Harris (ed.), *Cognitive processing in biinuals*, pp. 549–60. Amsterdam: North Holland Elsevier.

Paradis, M. (1990). Language lateralization in bilinguals: Enough already! *Brain and Language*, 39, 576–86.

Pienemann, M. (1988). Determining the influence of instruction on L2 speech processing. *AILA Review*, 5, 40–72.

Pienemann, M., (in press). *Language processing and second language development: Processability theory*. Amsterdam: John Benjamins.

Schwartz, B. D., & Tomaselli, A. (1990). Some implications from an analyisis of German word order. In W. Abraham, W. Kosmeijer and E. Reuland (eds.), *Issues in Germanic syntax*, pp. 251–74. New York: Mouton de Gruyter.

Vosse, T., & Kempen, G. (1991). A hybrid model of human sentence processing: Parsing right-branching, centre-embedded and cross-serial dependencies. *Proceedings 2nd Intenational Conference on Parsing Technologies*. Cancun, Mexico.

Weber-Fox, C. M., & Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience*, 8, 231–56.

https://doi.org/10.1017/S136672899800008X Published online by Cambridge University Press