Transfer and typological proximity in the context of second language processing
Gisela Håkansson, Manfred Pienemann and Susan Sayehli

Second Language Research 2002; 18; 250
DOI: 10.1191/0267658302sr206oa

The online version of this article can be found at:
http://slr.sagepub.com/cgi/content/abstract/18/3/250

Published by:

SAGE

http://www.sagepublications.com

Additional services and information for Second Language Research can be found at:

Email Alerts: http://slr.sagepub.com/cgi/alerts

Subscriptions: http://slr.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.co.uk/journalsPermissions.nav

Citations http://slr.sagepub.com/cgi/content/refs/18/3/250
Transfer and typological proximity in the context of second language processing

Gisela Håkansson University of Lund, Sweden, Manfred Pienemann University of Paderborn, Germany and Susan Sayehli University of Lund, Sweden

In this article, the issue of cross-linguistic influence in second language acquisition is examined from a processing perspective. Applying Processability Theory as the theoretical framework we claim that second language (L2) learners can only produce forms they are able to process. We thus argue that the first language (L1) influence on the L2 is developmentally moderated. Data were collected from German L2 learners with Swedish as their L1. Twenty informants participated in the study, 10 in their first year of German (13 years of age) and 10 in their second year of German (14 years of age). Both languages involved are typologically very close but not mutually intelligible. The results show that Swedish learners of German do not transfer the verb-second structure from their L1 to the L2 even though this structure is identical in both languages. Instead they start out with canonical word order and subsequently produce an intermediate structure (adv NP subj V X), which is ungrammatical in the L1 and the L2. These observations support the idea of a developmentally moderated transfer. The results clearly contradict the predictions from the ‘full transfer/full access' hypothesis (Schwartz and Sprouse, 1994; 1996).

I Introduction

In this article we will examine the issue of cross-linguistic influence in second language acquisition (SLA) from a processing perspective. The theoretical assumptions underlying this article are those made in Processability Theory (PT; Pienemann, 1998) which include, in particular, the following two hypotheses:

1) that first language (L1) transfer is constrained by the processability of the given structure; and

2) that the initial state of the second language (L2) does not necessarily equal the final state of the L1 (contrary to the assumption made by Schwartz and Sprouse, 1996), because there
is no guarantee that the given L1 structure is processable by the under-developed L2 parser.

In this article we present new empirical evidence supporting these claims. The evidence is derived from a study of the developing interlanguage syntax of learners whose L1 and L2 are typologically very close but not mutually intelligible. More specifically, the study focuses on the acquisition of German word order by native speakers of Swedish. These two languages share many regularities relating to word order. They therefore constitute a prime test case for arbitrating between the full transfer hypothesis (Schwartz and Sprouse, 1996) and the position of processing-constrained transfer advocated here. The logic of this test is quite straightforward: if the initial state of the L2 is the final state of the L1, then forms contained in both languages ought to appear early on in L2 acquisition. If, however, L1 transfer is constrained by processability, then forms higher up the processability hierarchy ought not to be transferred initially even if they are identical in both languages.

The key assumption of the processing perspective in SLA is that L2 learners can produce only those linguistic forms for which they have acquired the necessary processing prerequisites (Pienemann, 1998). Therefore PT predicts that, regardless of linguistic typology, only those linguistic forms that the learner can process can be transferred to the L2. These claims are operationalized in PT by being embedded in a coherent theoretical framework of L2 processing. To illustrate this operationalization, we give a summary of PT and characterize the lexical and hence language-specific nature of the processing of Swedish and German syntax within this framework.

The article is structured as follows; first the study of the acquisition of German by Swedes is reported. Then PT is summarized and utilized to explain the empirical findings of constraints on L1 transfer. Finally the concept of developmental constraints on transfer is juxtaposed with that of Schwartz and Sprouse’s (1996) full transfer hypothesis.

II The case of nontransfer from a typologically closely related language

1 Swedish and German word order

For the purpose of this article we focus on the position of the verb in Swedish and German main clauses. Both languages follow an SVO pattern in affirmative sentences if the grammatical subject is in initial position, as can be seen in (1) and (2).
1) Han reste till Stockholm.  (Swedish)
   ‘He went to Stockholm.’

2) Er fuhr nach Stockholm.  (German)
   ‘He went to Stockholm.’

When the sentence-initial position is occupied by a nonsubject, verb and subject are inverted in both languages, as shown in (3) and (4) (compare Pienemann, 1998; Pienemann and Håkansson, 1999).

3) Igår reste han till Stockholm.  (Swedish)
   ‘Yesterday he went to Stockholm.’

4) Gestern fuhr er nach Stockholm.  (German)
   ‘Yesterday he went to Stockholm.’

In other words, Swedish and German have at least three word-order rules in common: canonical word order (SVO), adverb fronting (ADV) and subject–verb inversion (INV). Taken together, these regularities (plus the fronting of other constituents) yield a constant verb-second position for all affirmative main clauses in both languages. This regularity is frequently referred to as ‘verb-second' or ‘V2' (compare Meisel, 1992; Holmberg and Platzack, 1995). Hence, both languages are V2-languages, in contrast to English. Apart from this, Swedish and German also contain other word-order rules which they do not necessarily have in common.

It is important to note in the present context that the grammatical rules the two languages have in common have been found to be acquired in the same sequence by L2 learners (compare Meisel et al., 1981; Pienemann, 1998; Pienemann and Håkansson, 1999), namely:

1st  SVO
2nd  ADV
3rd  INV

This implies that for an extensive period of time L2 learners of both languages produce grammatically unacceptable sentences containing ADV (or wh-fronting) but not INV. Examples for both L2s are given in (5) and (6).

5) * Igår han reste till Stockholm.  (Swedish)
   ‘Yesterday he went to Stockholm.’
6) * Gestern er reiste nach Stockholm.  
   (German)  
   ‘Yesterday he went to Stockholm.’

In both languages, learners can avoid the INV problem if they do not prepose the adverb. In that case the adverb would appear in final position. Depending on the class of the adverb, this yields either a grammatically acceptable or unacceptable sentence.

2 Constraints on L1 transfer

The assumption that L1 transfer may be developmentally constrained is not new in SLA research. Wode (1976; 1978) demonstrates such constraints for the acquisition of negation and interrogatives. He shows that German learners of English produce certain L1 forms only after they have developed the structural prerequisites in the L2. Zobl (1980) shows similar phenomena as does Kellerman (1983). What PT adds to the concept of developmental constraints on transfer is an explicit formal framework for specifying these constraints.

A clear trend towards constraints on transfer is also evident in a typological survey of L1 transfer. In Table 1 we have compiled a survey of studies of L2 word order, especially subject–verb inversion (or ‘V2’), in different typological constellations. Table 1 lists from left to right the basic word order of the L1, the L2 and the interlanguage (IL). It also states if the pattern found in the IL coincides with the L1 and might therefore superficially be considered a case of transfer. The overall pattern that emerges from this comparison is that there is a clear preference in early IL to follow canonical word order (SVO) without verb-second (SVO/–V2), irrespective of the word order in the source and target language. In other words, if the L1 does not have V2, then the IL will not have V2. However, the reverse is not true. If the L1 does have V2, the IL will nevertheless not contain V2. We argue that this is because V2 requires additional processing resources. This claim will be substantiated within PT.

The one crucial case not tested in the above typological scenario is a constellation in which V2 is contained in the L1 and the L2. If, in this constellation, V2 does not appear in the early IL, then this constitutes very strong evidence of developmental constraints on L1 transfer. Such a constellation exists in the case of Swedish and German. These languages share the +V2 property but are different enough to be true cases of SLA. They are not mutually comprehensible, as for instance Swedish and Danish are.
Given the identity of the above grammatical structures and their processing procedures in Swedish and German, it will be crucial to see what Swedish learners of German do in this typological scenario. If they do have access to all L1 knowledge and routines, as assumed by Schwartz and Sprouse (1994; 1996), then these learners ought to be able to transfer SVO, ADV and INV (or ‘V2’) as one package. In other words, one would expect two things:

1) INV ought to appear from the beginning of the acquisition process; and
2) the ungrammatical structure ‘ADV without INV' ought not to appear at all.

The study described below is designed to put this to an empirical test.
3 Research design

Data were collected from participants learning German as a foreign language in a Swedish secondary school (Sayehli, 2001). Twenty informants participated in the study, 10 in their first year of German (13 years of age) and 10 in their second year of German (14 years of age). All Swedish students in this age bracket have had English as a school subject for 2–3 years. German is therefore the informants' third language. We return to this point in a separate section at the end of this article. Specifically designed tasks were developed with the aim of eliciting narratives. The informants were interviewed by a native speaker of German. The interviewer had only a limited knowledge of Swedish, which helped motivate the informants to use German.

An anonymous reviewer was concerned that the informants in our study acquire German in a formal setting and that therefore ‘there is a good chance that some of these data represent learned knowledge rather than linguistic competence.’ In this context one should remember that one of the key features of this data collection is the time-constrained nature of the language production process. The learners produced the data in a conversational setting, thus being subject to the same constraints on word access and the computing of syntactic structures, etc. as any other speaker. Therefore, whatever they produce must be taken as evidence of their language processing skill and their underlying linguistic knowledge.

Presumably the reviewer had in mind a distinction between learned knowledge and linguistic competence. However, if learned knowledge results in language processing, how can it be differentiated empirically from linguistic competence? In fact, we believe that this distinction is a red herring in the context of this study. First, Paradis' (1994) analysis of the available neurophysiological evidence supports the view that explicit knowledge is not transformed into implicit knowledge. The distinction is therefore poorly motivated. Secondly, empirical studies of formal SLA (e.g., Pienemann, 1987; 1989) have demonstrated that there is in fact no difference between natural and formal acquisition in the order of L2 development.

4 Analysis and results

The oral L2 production of all informants was transcribed in full, and a distributional analysis was carried out. The results are shown in Tables 2 and 3. Both tables follow the same layout, which will be
read as an implicational table for the purpose of this article. The informants are listed vertically, and a count of the number of utterances and words is listed horizontally for each informant together with frequency counts of the total number of declarative main clauses and structures discussed above: (1) preposed adverbs (ADV) and (2) INV. All counts were carried out on the declarative main clauses produced by the informants.

Reading Table 2 as an implicational table, one can see that it supports a strict implicational ordering of the three structures concerned: SVO < ADV < INV. One observation is immediately striking in Table 2. None of the informants produced even one case of inversion, although the structure forms part of their native language, and it was contained in the L2 input. What is even more puzzling is the fact that 6 of the 10 informants produced ADV in the absence of INV, which is an ungrammatical structure, in both the L1 and the L2. This clearly falsifies Schwartz and Sprouse's (1994; 1996) ‘full transfer / full access' claim.

The same data-collection procedure and distributional analysis was carried out with learners of second-year German. The analysis is shown in Table 3. One general difference between the two groups of learners is that the second-year group produced ‘more language' than the first-year group. This is evident in the mean length of utterance which rises from 2.91 words per utterance in the first-year group to 3.24 in the second-year group. The other difference is the developmental status of the two sets of 10 L2 samples. The second-year group contains a greater number of advanced structures (ADV and INV) than the first-year group. In other words, the distributional analysis of the speech samples is congruent with the

Table 2  Results from first-year German, after 8 months of exposure

<table>
<thead>
<tr>
<th>Name</th>
<th>Utterances</th>
<th>Words</th>
<th>Declarative main clauses with subject and verb</th>
<th>Declarative main clauses with preposed adverbs</th>
<th>Declarative main clauses with inversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gelika</td>
<td>78</td>
<td>279</td>
<td>14</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Robin</td>
<td>125</td>
<td>308</td>
<td>14</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Emily</td>
<td>85</td>
<td>237</td>
<td>14</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Kennet</td>
<td>86</td>
<td>206</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Erika</td>
<td>92</td>
<td>284</td>
<td>15</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Sandra</td>
<td>95</td>
<td>276</td>
<td>16</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Johann</td>
<td>100</td>
<td>294</td>
<td>12</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Eduard</td>
<td>196</td>
<td>551</td>
<td>22</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Cecilia</td>
<td>117</td>
<td>295</td>
<td>21</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Anna</td>
<td>88</td>
<td>357</td>
<td>21</td>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>
simple and plausible expectation that the second-year group is
developmentally more advanced in the L2 than the first-year group.
Otherwise the second sample confirms the result of the first sample
with the same implicational pattern.

The most remarkable finding is that even after 16 months of
learning German only four students of the second-year group
produce INV, despite the fact that their L1 contains exactly the
same structure contained in the L2 input. Additionally, 7 out of the
10 learners in the second-year group produce a structure (+ADV
but –INV) that is ungrammatical in both their L1 and their L2. An
example is given in (7).

7) Cecilia: * Dann er waschen eh der Schlange.
then he wash  eh the snake

Given the cross-sectional nature of the data, it is permissible to pool
the data to obtain one comprehensive implicational table which is
presented in Table 4. Table 4 is laid out in a manner similar to Tables
2 and 3. One difference is that raw scores were converted into
binary notation using the emergence criterion (Meisel et al., 1981;
Pienemann, 1998) where + marks ‘acquired’ and – marks ‘not
acquired’.

As expected, the coefficient of scalability is 1.0 (Hatch and
Lazaraton, 1991). In other words, this database contains no
exceptions to the implicational pattern. In addition, Table 4
documents all stages of the implicational pattern with a minimum
of 4 speakers each. This adds to the strength of the observation.

Summing up, this database demonstrates that Swedish learners

<table>
<thead>
<tr>
<th>Name</th>
<th>Utterances</th>
<th>Words</th>
<th>Declarative main clauses with subject and verb</th>
<th>Declarative main clauses with preposed adverbs</th>
<th>Declarative main clauses with inversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Camilla</td>
<td>87</td>
<td>266</td>
<td>14</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mats</td>
<td>73</td>
<td>218</td>
<td>13</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ceci</td>
<td>177</td>
<td>564</td>
<td>39</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Mateus</td>
<td>101</td>
<td>376</td>
<td>30</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Peter</td>
<td>145</td>
<td>477</td>
<td>44</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Karolin</td>
<td>309</td>
<td>1003</td>
<td>66</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>Johan</td>
<td>103</td>
<td>368</td>
<td>29</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Zofie</td>
<td>120</td>
<td>347</td>
<td>24</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Zandra</td>
<td>127</td>
<td>365</td>
<td>14</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Caro</td>
<td>104</td>
<td>390</td>
<td>30</td>
<td>8</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 3: Second-year German, after 16 months of exposure.
Transfer and typological proximity of German, a verb-second language, do not transfer the verb-second structure from their L1 to the L2, even though the structure is identical in both languages and is contained in the learning input. Instead, they start out with canonical word order (NP_{subj} V NP_{obj} adv) and subsequently produce an intermediate structure (adv NP_{subj} V X) which is ungrammatical in the L1 and in the L2. All these observations clearly contradict the predictions to be derived from the ‘full transfer/full access’ hypothesis (Schwartz and Sprouse, 1994; 1996).

In the following section we explain the above observations on the basis of the processability of L2 structures, which constrains L1 transfer. For this purpose, the explanatory framework that is utilized here, Processability Theory, is summarized briefly below.

### III Developmental constraints on L1 transfer

#### 1 Processability Theory: a brief overview

Processability Theory (PT) (Pienemann, 1998) is based on a universal hierarchy of processing procedures which is derived from the general architecture of the language processor. This hierarchy is related to the requirements of the specific procedural skills needed for the target language. In this way, predictions can
be made for language development that can be tested empirically.

The view on language production followed in PT is largely that described by Levelt (1989). This 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) and on which the corresponding section of Levelt's model is based. The basic premises of that view are the following.

- **Premise 1:** Processing components, such as procedures to build NPs, etc., are relatively autonomous specialists that operate largely automatically. Levelt (1989) describes such grammatical procedures as ‘stupid', because their capacity is strictly limited to the very narrow but highly efficient handling of extremely specific processing tasks (e.g., NP-procedures, VP-procedures). The automaticity of these procedures implies that their execution is not normally subject to conscious control.

- **Premise 2:** Processing is incremental. This means that surface lexico-grammatical form is gradually constructed while conceptualization is still ongoing. One key implication of incremental language processing is the need for grammatical memory. For the next processor to be able to work on still-incomplete output of the current processor and for all of this to result in coherent surface forms some of the incomplete intermediate output has to be held in memory.

- **Premise 3:** The output of the processor is linear, while it may not be mapped onto the underlying meaning in a linear way. This is known as the ‘linearization problem' (Levelt, 1981) which applies to the mapping of conceptual structure onto linguistic form as well as to the generation of morpho-syntactic structures. One example is subject–verb agreement as illustrated in the sentence *She gives him a book.* The affixation of the agreement marker to the verb depends, amongst other things, on the storage of information about the grammatical subject (namely number and person), which is created before the verb is retrieved from the lexicon.

- **Premise 4:** Grammatical processing has access to a grammatical memory store. The need for a grammatical memory store derives from the linearization problem and the automatic and incremental nature of language generation. Levelt (1989) assumes that grammatical information is held temporarily in a grammatical memory store which is highly task-specific and in which specialized grammatical processors can deposit information of a specific nature (e.g., the value of diacritic features, such as the values for 'person', 'number', etc.).
Kempen and Hoenkamp's (1987) Incremental Procedural Grammar, the locus of the grammatical buffer is the specialized procedures which process NPs, VPs, etc. Pienemann (1998) presents evidence from on-line experiments and aphasia in support of these assumptions (e.g., Engelkamp, 1974; Cooper and Zurif, 1983; Paradis, 1994; Zurif et al., 1994).

The process of incremental language generation as envisaged by Kempen and Hoenkamp (1987) and Levelt (1989) is exemplified in Figure 1, which illustrates some of the key processes involved in the generation of the example sentence *A child gave the mother the cat*. The concepts underlying this sentence are produced in the Conceptualizer.

![Figure 1 Incremental language generation](http://slr.sagepub.com)
The conceptual material produced first activates the lemma \textit{CHILD} in the lexicon. This search is started from within the lexicalization system, a subsystem of the grammatical encoder. 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. Given certain conceptual specifications the lemma ‘A’ is activated and the NP-procedure attaches the branch Det to NP.

During this process the diacritic features of Det and N are checked against each other. This implies that the grammatical information ‘singular’ is extracted from each of the two lemmas at the time of their activation and is then stored in NP until the head of the phrase is produced. This process of exchange of grammatical information is a key feature of language production. Below we utilize Lexical Functional Grammar (LFG), which has the capacity to model the exchange of grammatical information by feature unification.

The production process has now 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 the establishment of 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.

While the above process was still ongoing, the next conceptual fragment would have been processed in parallel and the output of the Formulator\textsuperscript{1} would have been delivered to the Articulator. This means that new conceptualization occurs while the conceptual structure of the previous iteration is being produced. The whole process then moves on from iteration to iteration.

Kempen and Hoenkamp’s research implies that in the process of incremental language generation the following processing procedures and routines are activated in the following sequence:

1) lemma access;
2) the category procedure;
3) the phrasal procedure;
4) the S-procedure;
5) the subordinate clause procedure, if applicable.

\textsuperscript{1} In Levelt’s (1989) model the Grammatical Encoder and the Phonological Encoder make up the Formulator.
Pienemann (1998) hypothesizes that these key grammatical encoding procedures are arranged according to their sequence of activation in the language generation process, and that this sequence follows an implicational pattern in which each procedure is a necessary prerequisite for the following procedures. The basic thesis of PT is that in the acquisition of language processing procedures the assembly of the component parts will follow the above-mentioned implicational sequence. The key to predicting which grammatical structures are processable, and in which sequence, lies with the question of which pieces of grammatical information can be exchanged between which constituents, given the availability of the different procedures and their storage capacity.

Pienemann (1998: 73) points out that the above processing procedures are fully operational only in native speakers of a language, not in 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. In this context it is important to ensure that Levelt's model (and Kempen and Hoenkamp's specific section of it) can, in principle, account for language processing in bilinguals, since second language acquisition will lead to a bilingual language processor. Pienemann (1998: 73)

PT utilizes, amongst other things, de Bot's (1992) work to apply the processability hierarchy to bilingual language production. 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. The key assumption of de Bot's work for L2 processing is that in all cases where the L2 is not closely related to the L1, different (language-specific) procedures have to be assumed. Pienemmann (1998: 78–79) therefore concludes that most of the above processing procedures have to be acquired by the L2 learner. He cites diacritic features such as 'tense', 'number', 'gender' and 'case', which vary between languages, as obvious examples of cross-linguistic differences in the lexical prerequisites for language processing.

When a given learner has not developed a specific element of the implicational hierarchy of processing procedures assumed by PT, the hierarchy will be cut off in the procedural grammar of the
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. In other words, it is hypothesized by PT that processing procedures and the capacity for the exchange of grammatical information will be acquired in their implicational sequence as depicted in Table 5 where $t_1$, $t_2$, etc. refer to different points in the course of language development.

Pienemann (1998) implemented this hierarchy in LFG, because LFG has been shown to be compatible with Levelt's overall design of the language generator. In this way, developmental patterns can be related to an overall hierarchy of processability, and they can be compared across languages.

For reasons of space we do not demonstrate this implementation process for Swedish and German and refer the reader to Pienemann (1998) for an extensive treatment of the issue. For the purpose of this article it suffices to provide a simplified summary of the implementation process and the resulting processing hierarchy applied to Swedish and German in the form of Tables 6 and 7.

The reader should note that the acquisition sequences of Swedish and German word order that we referred to above can be identified in these tables and that they follow from the overall hierarchy of processability.

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Hypothetical hierarchy of processing procedures</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_1$</td>
<td>$t_2$</td>
</tr>
<tr>
<td>S’-procedure (EmbeddedS)</td>
<td>–</td>
</tr>
<tr>
<td>S-procedure</td>
<td>–</td>
</tr>
<tr>
<td>Phrasal procedure (head)</td>
<td>–</td>
</tr>
<tr>
<td>Category procedure (lexical category)</td>
<td>–</td>
</tr>
<tr>
<td>Word/lemma</td>
<td>+</td>
</tr>
</tbody>
</table>

Note: $t_1$, $t_2$, etc. refer to different points in the course of language development.
Source: Pienemann, 1998: 79
Developmentally moderated transfer

The crucial point to be explained in the present discussion of developmental constraints on L1 transfer is why Swedish learners of German do not transfer INV from their L1 to the L2. The point implied by PT is that with the given architecture of the language generator there is no guarantee that one can simply utilize L1 procedures for the L2. Pienemann (1998: 80–81) argues that such a 'bulk transfer' would lead to internal problems:

> because all of the above processing procedures need to be orchestrated in a language-specific way. If any one of them is missing or incompatible with the rest, the Formulator is inoperable. If, for instance, the lexical category information is missing, category and phrasal procedures cannot be called. If diacritic features are missing or have no values or values which are not compatible with those listed in agreeing phrases or if they are incompatible with the Functorisation Rules, then the processor will be inoperable. (Pienemann, 1998: 80)

PT does not imply, however, that the learner will never attempt to form diacritic features and Functorization Rules that reflect L1 regularities. Instead, the theory does imply processing constraints on L1 transfer:

> a 'bulk-transfer' of the L1 Formulator would lead to very unwieldy hypotheses.

German learners of English, for instance, would have to invent large sets of
diacritic features for nouns, verbs and adjectives without any evidence of their existence in the L2, since German definite determiners express a complex set of diacritic features of the noun (three genders and two numbers). Since English nouns do not contain these diacritic features in nouns, the complex system of definite determiners presented in Table 8 corresponds to merely one English grammatical morpheme (‘the’).

In this case the simplest structural solution would be to abandon the L1 diacritic features altogether. This would in fact reproduce a situation which is close to the English determiner system. However, the relationship between L1 and L2 diacritic features may be more complex than in the above example, with two intersecting sets of diacritic features and different form–function relationships in L1 and L2. In other words, there is potentially a multitude of L1 features only some of which are applicable to the L2. (Pienemann, 1998: 81)

In essence the lack of psychological plausibility present in the ‘bulk transfer’ approach forms a logical argument in favour of processing constraints on L1 transfer, the position assumed by PT:

I hypothesise that the L1 Formulator will not be ‘bulk-transferred’. Instead, the learner will re-construct the Formulator of the L2. This would not exclude that in the course of this process L1 procedures be utilised. However, I hypothesise that such L1 transfer always occurs as part of the overall reconstruction process. (Pienemann, 1998: 81-82)

The case of constraints on the transfer of morphological and lexical regularities is relatively straightforward as the example of the determiner illustrates. It may be less obvious why the same constraints should also be operational on word order, particularly in the case of two languages that share several identical features. To appreciate this point it may be useful to consider the lexical nature of word order regularities within an LFG framework and indeed in Levelt’s model of language generation.

This point has been illustrated by Pienemann (1998: 99–102) with the acquisition of German separable verbs. This has been described as a gradual lexical acquisition process that is based on a number of alterations of the initial c-structure rule R1 for canonical word
order when it is modified in later development into R2. One alteration concerns the introduction of VP as a constituent, which is necessary to account for a range of phenomena in German, as we 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 $(\text{NP}_{\text{obj}1})(\text{NP}_{\text{obj}2}) \text{ V}$. This ensures that V will only occur in second position unless ‘licensed’ by a V that takes V-COMP.

\[(\text{R1}) \text{ S } \rightarrow \text{ NP}_{\text{subj}} \text{ V } (\text{NP}_{\text{obj}1})(\text{NP}_{\text{obj}2})\]

\[(\text{R2}) \text{ S } \rightarrow \text{ NP}_{\text{subj}} \text{ VP} \]

\[\text{VP } \rightarrow \text{ V } \left( (\text{NP}_{\text{obj}1})(\text{NP}_{\text{obj}2}) \right) \]

\[\text{V-COMP } \rightarrow (\text{NP}_{\text{obj}1})(\text{NP}_{\text{obj}2}) \text{ V} \]

Apart from this change in c-structure rules, Pienemann (1998) assumes that the learner gradually re-analyses the verbs of his or her IL, by analysing AUX and V as two separate entries and by adding the feature AUX to the lexical features of V.

The relevant part of that analysis is as follows:

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). Let us take sentence (8) as an example:

8) er hat ein Bier getrunken
   he has a beer drunk
   ‘He has drunk/drank a beer.’

The c-structure of (8) is represented in the form of a simplified tree diagram in (9). The simplified lexical entries for the verbs in (10) are as shown in Figure 2.
This set of entries and rules, etc. ensure two things which are of relevance here: (1) that a particular (at this stage not necessarily the correct) morphological form of the lexical 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.

(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. (Pienemann, 1998: 100–102)

This unification process is the key point in our discussion of L1 transfer. The above example illustrates that word order phenomena depend crucially on the correct annotation of lexical entries, which differ in their diacritic features even between related languages.

Cross-linguistic differences in diacritic features are further illustrated by the comparison of two lexical classes and their diacritic features in closely related languages, which is shown in Table 9. In other words, the procedures that achieve the different positions are highly language specific, and the lexical information involved is not necessarily compatible even between two closely related languages.
In the case of the Swedish–German contrast there is also a host of syntactic phenomena which apply to one of the languages only; for instance, a complex array of constraints on the placement of negation that apply to Swedish only (compare Pienemann and Häkansson, 1999). The Swedish learner of German is therefore faced with the following situation: the phonology of the two languages is markedly different, the lexical items are not mutually intelligible, and the structure of the lexicon is different despite certain overlaps. A similar overlap also applies to syntax.

In this situation the learner has the task of discovering which lexical features and which syntactic patterns are shared by the two languages. Pienemann (1998) points out that the existing similarities between the languages are not obvious to the learner, because they are distributed amongst a host of typological differences:

While it may be clear to the linguistic analyst which of the diacritic features of the L1 apply to the L2 and which syntactic patterns are shared between the languages, there is no obvious a priori way for the learner to know this. A random choice of features would be likely to generate procedures which are incompatible with the rest of the Formulator. Unless the learner simply limits herself or himself to the L1 Formulator, thus re-lexifying the L1 (as may be possible in mutually intelligible languages) and not acquiring the L2, there is no other obvious choice than to re-construct the set of diacritic features and syntactic routines specific to the L2. (Pienemann, 1998: 81)

In other words, according to PT, the reconstruction of the L2 and developmental constraints on L1 transfer follow from the hierarchical nature of the learning task. In this scenario there is no other logical point of departure for this reconstruction process than the beginning of the processability hierarchy because the latter is stripped of all language-specific lexical features and syntactic routines. It would therefore be logical for this reconstruction
process to follow the path described in the processability hierarchy and for L1 knowledge and skills to become accessible once they are processable in the developing system.

Summing up, PT implies the hypothesis that the L1 Formulator will not be ‘bulk-transferred’, because the processing of syntax is lexically driven and the processor relies on highly language-specific lexical features. Instead, the learner will reconstruct the Formulator of the L2. This would not exclude the possibility that in the course of this process L1 procedures will be utilized. However, it is hypothesized that such cases of L1 transfer occur as part of the overall reconstruction process. This means that L1 transfer is developmentally moderated and will occur when the structure to be transferred is processable within the developing L2 system.

3 The ‘English illness’

Given that in our study German was in fact the third language of the informants and that English was the second, it may be easy to conclude that the nonapplication of INV (or V2) was due to transfer from English. In fact, this explanation is popular amongst Swedish school teachers of German and has also been suggested by Ruin (1996) and Naumann (1997). Swedish teachers of German disrespectfully term this phenomenon the ‘English illness’.

However, such a proposal is not compatible with the data from our study. First of all one would need to consider how the transfer-from-L2 hypothesis would be testable. Logically, the hypothesis would have to predict that all L2 word order constraints would be transferred, or at least all those that are shared by the L1, the L2 and the L3. Otherwise the transfer hypothesis would have no predictive power and could not be falsified, unless one added a separate theory predicting which items are to be transferred and which are not.

In the absence of such a theory we will test the transfer-all hypothesis. To follow our argument, it is important to remember that the data from the above study show a strictly implicational development. It is evident from this analysis that 6 of the 20 learners produce SVO only and no ADV. If one followed the transfer view, they would appear to have transferred selectively only one word order pattern known from their L2 (English). This clearly falsifies the transfer-all hypothesis and leaves the selective-transfer hypothesis with the problem of making no testable prediction as to when transfer will take place.

We mentioned above that our study falsifies predictions made by
the ‘full transfer/full access hypothesis’ (Schwartz and Sprouse, 1994; 1996) according to which the learner begins with the L1 grammar as the initial state of the L2 grammar. With this assumption the Swedish learners of German in our study would be predicted to use the target V2 word order from the very beginning. As we demonstrated above, this prediction is incorrect.

Given this blunt falsification of the ‘full transfer/full access hypothesis’ it may be worthwhile examining the original empirical basis of this hypothesis. Schwartz and Sprouse base their argument on data from one learner, Cevdet, a Turkish learner of German. The data for their Stage 1 originates from an interview made when Cevdet had already spent one year in Germany. Furthermore, he had received formal German instruction over a period of 6 months for 10 hours per week.

This amount of exposure and formal language learning makes it highly probable that ‘Stage 1' is indeed not the initial stage in Cevdet's acquisition of German and that he had instead passed through several stages of acquisition when this sample was collected. Therefore it is rather doubtful if these data are relevant for claims on the initial state or on transfer in the initial stage of acquisition.

By comparison, the Swedish data consist of 20 samples instead of one as in Schwartz and Sprouse's study. The informants had only formal exposure to the L2. The less advanced group had had a total of 50 hours of exposure, compared to 240 hours of formal exposure alone in the case of Cevdet. The amount of exposure in the L2 environment can only be guessed. It is therefore very likely that the Swedish sample covers a far earlier segment of L2 development. This is also supported by the perfect consistency in the implicational analysis of the data. This constitutes the strongest possible evidence that the data represent a continuous developmental pattern starting from canonical word order and following a developmental sequence that is predicted by the processability hierarchy.

IV References


du Plessis, J, Solin, D, Travis L. and White, L. 1987: UG or not UG, that is the question. Second Language Research 3, 56–75.


