Constraining second language word order optionality: scrambling in advanced English-German and Japanese-German interlanguage

Holger Hopp

Second Language Research 21,1 (2005); pp. 34–71

Constraining second language word order optionality: scrambling in advanced English–German and Japanese–German interlanguage

Holger Hopp *University of Groningen*

This study documents knowledge of UG-mediated aspects of optionality in word order in the second language (L2) German of advanced English and Japanese speakers ($n = 39$). A bimodal grammaticality judgement task, which controlled for context and intonation, was administered to probe judgements on a set of scrambling, topicalization and remnant movement constructions. Given first language (L1) differences and Poverty of the Stimulus, English and Japanese learners face distinct learnability challenges. Assuming Minimalist grammatical architecture (Chomsky, 1995), convergence on the target language would entail the unimpaired availability of Universal Grammar (UG), i.e., computational principles and functional features beyond their L1 instantiation. Irrespective of L1, the L2 groups are found to establish systematic native-like relative distinctions. In addition, L1 transfer effects are attested for judgements on scrambling. It is argued that these findings imply that interlanguage grammars are fully UG constrained, whilst initially informed by L1 properties.

I Introduction

The central question in generative nonnative or second language (L2) acquisition research is whether the domain-specific, innate restrictions on the hypothesis space explored in first language (L1) acquisition as given by Universal Grammar (UG) equally delimit the range of options in adult interlanguage (IL). Some researchers argue that post-pubescent

Address for correspondence: Holger Hopp, Department of General Linguistics, University of Groningen, Postbus 716, 9700 AS Groningen, The Netherlands; email: hopp@let.rug.nl

 10.1191/0267658305sr246oa

L2 acquisition falls outside the bounds of UG (Bley-Vroman, 1997; Meisel, 1997; 2000). Within the generative paradigm, approaches diverge on whether parametric restructuring is fully available (e.g., Epstein *et al.*, 1996; Schwartz and Sprouse, 1996), locally delimited to specific functional features (Hawkins and Chan, 1997) or globally impaired (Clahsen and Muysken, 1989; Meisel, 1997).

As recently argued by Schwartz and Sprouse (2000), decisive evidence about the nature of UG involvement in L2 acquisition comes solely from studies of the logical problem of language acquisition (Chomsky, 1986). In L2 acquisition, this obtains if (1) the target-language (TL) input underdetermines a UG-specified grammatical phenomenon and (2) the phenomenon is not instantiated in the learner's L1 (e.g., White, 2003). If in such cases L2 behaviour is target-like, logic dictates that IL grammars must be supported by domain-specific representations as given by UG.

In the Minimalist Program (Chomsky, 1995), typological diversity reduces to variation in the lexicon, whose items and features are concatenated by a universal derivational algorithm (C_{HL}) to output as well-formed Phonological Form (PF) and Logical Form (LF) representations. Overt syntactic movement is motivated by strong uninterpretable features that require checking against matching probes prior to Spell-Out. Whether and where constituents move is subject to language-particular parametrization of the functional lexicon, Lex_{FF}, where types of features and their strength are encoded. Minimalist grammatical architecture imposes a restrictive perspective on the language faculty, and its reflexes should be manifested in ILs if UG fully constrains L2 grammars.

The present article empirically investigates optional movement phenomena in German that are subject to a universal constraint, the Principle of Unambiguous Domination (UD; Müller, 1996; 1998). Since this UG principle refers to types of functional features, its operation is evidenced only in languages that instantiate interacting feature-driven A'-movements, e.g., scrambling and topicalization in German and Japanese. The grammar of English does not provide any overt evidence for the operation of UD, nor does the input available to learners of German. Building on previous work by Schreiber and Sprouse (1998), I report an empirical cross-linguistic study which shows that both advanced English–German and Japanese–German IL manifest the

precise distinctions enforced by the interaction of a language-specific functional lexicon and invariant computational principles, despite Poverty of the Stimulus. After demonstrating the robustness of grammaticality distinctions across L1s, proficiency levels, syntactic paradigms and individuals, I argue that these results require an organization of IL systems as envisaged by (Minimalist) generative theory.

This article is organized as follows: Section II discusses word-order optionality in German, English and Japanese in the framework of Minimalist syntax. In Section III, I argue that the particular constraints on word-order optionality in German are inducible neither from the properties of English grammar nor from the TL input. Based on these assumptions, Section IV outlines three positions on the role of UG in L2 acquisition and develops their predictions for the acquisition of word-order optionality in German. Section V presents the study and its findings. After discussing alternative accounts of the results in Section VI, I conclude that the findings implicate the full involvement of UG in adult L2 acquisition.

II Word-order optionality in German, Japanese and English

1 Scrambling

Unlike English with a relatively fixed subject–verb–object order (1), semi-free word-order languages, like German and Japanese, allow for scrambling, i.e., the optional linear reordering of verbal arguments.[1] Examples (2b) and (3b) illustrate scrambling in the complement domain, where the order of objects alternates. Examples (2c) and (3c) exemplify scrambling of objects across the subject.

1) a. John gave Mary the book.
   b. * John gave the book Mary.
   c. * I think that the book John gave (to) Mary.
2) a. … dass John Maria das Buch gab.
      … that John Mary the book gave
      'that John gave Mary the book.'
   b. … dass John [das Buch]$_1$ Maria t_1 gab.
      that John the book Mary gave
   c. … dass [das Buch]$_1$ John Maria t_1 gab.
      … that the book John Mary gave

[1]In generative analyses of German, the verb-final order in embedded clauses is generally taken to be the base order from which main clause verb-second is derived by movement (e.g., Vikner and Schwartz, 1996). In Japanese, an absolute head-final language, the position of the verb never changes.

3) a. John-ga Mary-ni sono hon-o watasita.
John-NOM Mary-DAT that book-ACC gave
b. John-ga [sono hon-o]$_1$ Mary-ni t_1 watasita.
John-NOM that book-ACC Mary-DAT gave
c. [sono hon-o]$_1$ John-ga Mary-ni t_1 watasita.
that book-ACC John-NOM Mary-DAT gave

Most approaches consider flexible argument ordering to be the formation of A'-chains with the surface order being derived from the base order by antecedent-trace relations. In Minimalist analyses, syntactic displacement is driven by strong uninterpretable functional features to be checked against their matching probes (Chomsky, 1995). Accordingly, whether a language instantiates scrambling reduces to the availability of a strong uninterpretable scrambling feature [scr] in the language-particular functional lexicon, Lex_{FF} (Oka, 1996; Sauerland, 1999). Scrambling typically occurs in particular information-structural and semantic contexts (e.g., Ishihara, 2001; Müller, 1999). In these contexts, a scrambling feature can optionally occur in the feature matrix of a maximal category (e.g., DP, PP, CP). In these cases, the [scr] feature requires checking against its probe prior to Spell-Out, and hence occasions overt movement of this XP. Scrambling features are checked against their matching probes that are realized in the heads of agreement projections (Grewendorf and Sabel, 1999). Analogously, topicalization and *wh*-movement are motivated by strong features on XPs, [top] and [wh], to be checked against their respective probes (C^0) before Spell-Out (Müller and Sternefeld, 1993).

On these assumptions, the cross-linguistic differences in (1) to (3) result from distinct repertoires of functional features and checking sites. German and Japanese license a scrambling feature, and $AgrO^0$ and $AgrS^0$ are checking heads for scrambling.[2] Since English does not license any kind of scrambling (1), the English functional lexicon can be taken not to supply a scrambling feature [scr] which could optionally be part of a numeration. For topicalization and *wh*-movement, all languages pattern similarly in that they all provide the same array of features, i.e., [top] and [wh]. Unlike in English, topicalized phrases are morphologically marked with a *wa*-suffix in Japanese (4a). Example (4b)

[2]In contrast to German, Japanese also licenses long scrambling out of finite clauses.

illustrates *wh*-movement in Japanese main clauses bearing the overt question marker *no* (Takahashi, 1993).

4) a. Konpyuutaa-wa Taroo-ga katta
Computer-TOP Taro-NOM bought
'A computer, Taro bought.'
b. Nani-o_1 John-wa [$_{CP}$Mary-ga t_1 tabeta ka] siritagetteru no
What-ACC John-TOP Mary-NOM ate Q want-to-know Q
'What does John want to know whether Mary ate?'

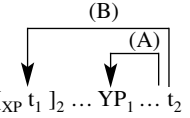
2 Remnant movement

Remnant movement describes movement of a maximal category across an element that has previously been extracted from it (e.g., den Besten and Webelhuth, 1990). This is illustrated in (5):

A) YP first moves out of the larger phrase, XP, leaving what remains of XP, i.e., the remnant;
B) then, the remnant phrase (subscripted 2) fronts across YP and bears an unbound trace.

Such remnant movement constellations are subject to a putatively universal constraint, formulated in terms of feature-driven movement types (6).

5)



6) Principle of Unambiguous Domination (UD) (Müller, 1996; 1998; Takano, 2000): In a derivation yielding the configuration … $[_A$ ... t_i ...$]_j$ ... B_i ... t_j ..., movement of B and movement of A may not be of the same type.

As stated in (6), UD prohibits remnant movement if both movements are of the same type, where type is defined in terms of features, e.g., [scr], [top] or [wh]. In scrambling languages, UD explains an unexpected asymmetry in movement options that cross-cuts the traditional A vs. A'-dichotomy. In the German (7), there is a sharp contrast in grammaticality between the same linear order of constituents depending on whether it is derived by topicalization or scrambling. In (7a), the object DP *den Wagen* ('the car') first scrambles over the adverb *schon* ('already'); then, the infinitival remnant *zu reparieren* ('to repair') topicalizes to sentence-initial position. The resulting order is grammatical, since the two movements are of different types, i.e., scrambling and

topicalization. In (7b), the same linear order is illicit in embedded clauses where scrambling of the object DP *den Wagen* feeds scrambling of the infinitival remnant *zu reparieren* (Grewendorf and Sabel, 1994).

7) a. $[_{\alpha}\, t_1$ Zu reparieren$]_2$ hat Peter $[_{DP}$ den Wagen$]_1$ schon t_2 versucht.
To repair has Peter the car already tried.
b. * Ich glaube, dass $[_{\alpha}\, t_1$ zu reparieren$]_2$ Peter $[_{DP}$ den Wagen$]_1$ schon t_2 versucht hat.
I think that to repair Peter the car already tried has.
'I think that Peter already tried to repair the car.'

The same contrast obtains in Japanese (*wh*)-scrambling (Koizumi, 1995). Remnant movement in cases where scrambling feeds *wh*-movement (8a) is licit, since different types of features, [scr] and [wh], are involved. By contrast, remnant scrambling across scrambling (8b) is disallowed.

8) a. $[_{CP}\, [_{SC}\, t_1$ donna-ni kirei-ni$]_2$ $[_{TP}$ biyoosi-ga $[_{DP}$ Mary-o$]_1$ t_2 sita] no]
how-much beautiful beautician-NOM Mary-ACC did Q
'How beautiful did the beautician make Mary?'
b. * $[_{TP}\, [_{SC}\, t_1$ kirei-ni$]_2$ $[_{TP}$ biyoosi-ga $[_{DP}$ Mary-o$]_1$ t_2 sita]]
beautiful beautician-NOM Mary-ACC did
'The beautician made Mary beautiful.' (Tsujioka, 2001: 492)

Importantly, English does not show the same grammaticality contrasts. As English does not have scrambling, an A'-movement, English does not allow for any grammatical remnant movements. Since topicalization independently creates strong islands (Rochemont, 1989), the interaction of topicalization and *wh*-movement cannot yield grammatical derivations (9a). In (9b), a case of remnant *wh*-movement across *wh*-movement (Lasnik and Saito, 1992), a violation of subjacency further adds to the UD violation.

9) a. * Who_1 do you think [the book]$_2$ she gave t_1 t_2?
b. * $[_{NP}$ Which picture of $t_1]_2$ do you wonder $[_{CP}$ who_1 she likes $t_2]$?

In sum, remnant movements are realized differently in the three languages due to the parametrization of the language-particular functional lexicon of formal features. Although a universal principle, UD discriminates between grammatical and ungrammatical remnant movements only in German and Japanese (Müller, 1998). As English does not license scrambling, English does not instantiate a grammaticality contrast in remnant movement configurations. Accordingly, there is no overt evidence of the operation of UD for A'-movement in English, since remnant A'-movement is always ungrammatical in English.

These typological asymmetries between scrambling and nonscrambling languages have different consequences for the learnability of word-order optionality for L2 acquirers of German.

III German word-order optionality and learnability

1 Word order variation and input

Acquiring optionality in word order constitutes a daunting learnability challenge, since evidence of the range and restrictions of word-order variation is likely to be sparse and ambiguous in the input. Learners need to establish whether any given construction is obligatory or allows for optional variation, and whether possible variants are truly optional or restricted to particular (interpretive) contexts (Papp, 2000).

In this section, I argue that the particular set of linearization options and restrictions offered in scrambling languages like German cannot be learnt on the basis of the available input or instruction alone (see also Schreiber and Sprouse, 1998). Consider the German word-order paradigm in (10) in this respect. Example (10) non-exhaustively charts movement options (scrambling and topicalization) in the context of so-called coherent infinitives, a class of verbs allowing extraction (Müller, 1998). I refer to the set of the sentences in (10) as the 'infinitival paradigm'. A homologous paradigm for PP extraction from DPs ('the DP paradigm') is introduced later.

10) Infinitival paradigm[3]

a. *Scrambling of complete phrase:*
Ich glaube, dass [den Wagen zu reparieren]$_1$ Peter schon t$_1$ versucht hat.
I think that the car to repair Peter already tried has
'I think that Peter already tried to repair the car.'

b. *Topicalization of complete phrase:*
[Den Wagen zu reparieren]$_1$ hat Peter schon t$_1$ versucht.
The car to repair has Peter already tried

c. *Remnant topicalization across scrambled phrase:*
[t$_1$ Zu reparieren]$_2$ hat Peter [den Wagen]$_1$ schon t$_2$ versucht.
To repair has Peter the car already tried

d. *Remnant topicalization (across scrambled phrase) across finite clause boundary:*
[t$_1$ Zu reparieren]$_2$ glaube ich [t′$_2$ hat Peter [den Wagen]$_1$ schon t$_2$ versucht].
To repair think I has Peter the car already tried

[3]As the various truth-conditionally synonymous derivations in (10) do not have equivalents in English, I do not gloss the differences between them in terms of information structure.

e_1. *Remnant scrambling across short-scrambled phrase:*
* Ich glaube, dass $[t_1$ zu reparieren$]_2$ Peter [den Wagen]$_1$ schon t_2 versucht hat.
I think that to repair Peter the car already tried has

e_2. *Remnant scrambling across medium-scrambled phrase:*
* Ich glaube, dass $[t_1$ zu reparieren$]_2$ [den Wagen]$_1$ Peter schon t_2 versucht hat.
I think that to repair the car Peter already tried has

f. *Remnant topicalization across topicalized phrase:*
* $[t_1$ Zu reparieren$]_2$ glaube ich [den Wagen]$_1$ hat Peter schon t_2 versucht.
To repair think I the car has Peter already tried

Example (10a) illustrates scrambling of complete phrases, in this case movement of the predicate of the verb *versuchen* ('to try') across the subject in the embedded clause. Example (10b) shows that the complete phrase can equally topicalize in main clauses, inducing verb second. The remaining sentences in (10) display remnant movement constellations. Example (10c) illustrates remnant topicalization across the DP *den Wagen* ('the car'), which has previously been scrambled. Example (10d) is another example of remnant topicalization, where the remnant fronts across a finite clause boundary. In both (10c) and (10d), the two instances of movement are of different types, i.e., scrambling and topicalization. Accordingly, these sentences are licit as per UD (6). Examples ($10e_1$), ($10e_2$) and (10f) are violations of UD. In ($10e_1$) and ($10e_2$), UD blocks remnant scrambling across short scrambling in the complement domain ($10e_1$) and across medium scrambling, i.e., across the subject ($10e_2$). Finally, (10f) marks an illicit case of remnant topicalization across a topicalized object in an embedded clause.[4]

Are the exact grammaticality contrasts in (10) learnable from the TL input? Let us consider two strategies for inferring grammaticality from the input (Bley-Vroman, 1997; Meisel, 1997): differences in the statistical frequency of occurrence of the variants in (10), and differences in surface word order patterns.

As for frequency, whilst learners might occasionally be exposed to some of the licit sentence types in (10) in naturalistic conversation, such selective encounters of particular optional variants by no means ensure convergence on the precise matrix of grammaticality contrasts. In general, inductive learning from the input relies on frequency and

[4]Note, however, that remnant topicalization across topicalization is independently ruled out by the fact that topicalization creates strong islands in the Germanic languages (e.g., Rochemont, 1989). Therefore, the effects of UD are somewhat blurred in (10f).

salience (e.g., Ellis, 2002). In other words, the input must robustly reflect the complete matrix of movement options, such that a gap in the paradigm can be identified. However, corpus studies demonstrate that the noncanonical word orders in (10), in particular scrambling of complex XPs and remnant movement, are highly infrequent in spoken and written German (Hoberg 1981; Schlesewsky *et al.*, 2000; Bornkessel *et al.*, 2002). The relative statistical difference between infrequent sentences and non-occurring ungrammatical sentences is thus very small. Therefore, observing the relative discourse frequency of noncanonical orders is unlikely to lead to a reliable distinction between rare licit and non-instantiated illicit sentences.

However, even if the frequency differences between the low-frequent and the ungrammatical variants were sufficiently robust, they could not be used as a reliable determinant of grammatical status. Due to the discourse-optionality of scrambling and topicalization, a learner cannot conclude from the absence of a possible ordering variant in the input that it is ungrammatical. This is so because inferring the ungrammaticality of a construction from its non-occurrence in the input (indirect negative evidence; Chomsky, 1981: 9) depends on the existence of contexts in which the given construction should obligatorily be used but fails to occur (e.g., Pinker, 1984). Only in such circumstances can noticing the absence of a string reliably disabuse the learner of its grammaticality. With respect to the sentences in (10), however, such a constellation never arises. The unmarked base order can serve as a pragmatically felicitous expression in any possible discourse context, given appropriate stress assignment (Höhle, 1982). In other words, there is no discourse context that requires a scrambled or topicalized sentence to be used instead of the base order. Consequently, even the persistent non-occurrence of noncanonical variants in the TL input could not be a reliable indicator of their ungrammaticality, since this non-occurrence might simply reflect the preferences of speakers rather than the grammaticality of sentences.

Let us consider the second option for inducing the grammaticality contrasts in (10) from observable surface properties of the input. Suppose that the full paradigm in (10), including asterisks for ungrammaticality, were open to inspection. Assume further with e.g., Meisel (1997) that, in their analysis of linguistic input, L2 learners

primarily refer to surface string orders. Such low-level distributional analyses of linear surface patterns would, however, fail to yield a descriptively correct generalization of the grammaticality contrasts in (10). Arguably, a learner is most likely to notice in (10) that the relative positions of the embedded transitive verb *zu reparieren* ('to repair') and its complement *den Wagen* ('the car') vary. Yet, extrapolating the grammaticality contrasts in (10) by positing a fixed (relative) surface order of the transitive verb *zu reparieren* and its complement [OBJ V] runs aground in light of the grammatical (10c–d) in which the verb precedes its object. However, allowing for the [V SUBJ OBJ] surface order of (10c–d) fails to rule out the ungrammatical ($10e_1$), which shows the same linear sequence. Similarly, the grammaticality contrasts cross-cut the difference between matrix and embedded clauses; compare, e.g., (10d) and (10f). It appears that cognitively plausible approaches towards inducing the constraint pattern in (10) in terms of the linear distribution of constituents could not produce the correct generalization even in the hypothetical case where the entire paradigm in (10) is unambiguously available to the learner. In sum, neither indirect negative evidence nor positive evidence exists in the input to prevent the learner from extending word-order freedom in German signalled by remnant topicalization to remnant scrambling.

Finally, neither explicit instruction nor negative evidence about the ungrammaticality of remnant scrambling across scrambling is systematically available. Remnant movement is neither treated in standard textbooks or English-language reference grammars of German, nor were the instructors of German we consulted consciously aware of the grammaticality contrasts in (10). Thus, the TL input and meta-linguistic instruction underdetermine the grammaticality distinctions governed by UD. Given that scrambling and topicalization as well as remnant movement *per se* are possible in German, one might expect learners to treat remnant scrambling equivalently.

L1 acquirers of German face exactly this learnability challenge in the face of input that underdetermines the ungrammaticality of certain remnant movements. Assuming Poverty of the Stimulus, domain-specific knowledge of the kind envisaged by UG is arguably essential to ensure that German speakers converge on a grammar bearing the

distinctions outlined in Section II and in (10). On Minimalist assumptions, these asymmetries follow from the interaction of invariant computational principles (UD) and the language-particular repertoire of formal features ([scr], [top]).

2 L2 learnability and cross-linguistic differences

In the Minimalist Program, cross-linguistic differences do not affect the computational principles of syntax. In the context of L2 acquisition, Hale (1996) therefore argues that the distinction between knowledge derived from the L1 and knowledge drawn directly from UG becomes impossible to test empirically in studies of L2 acquisition because the fundamental properties of any L2 grammar are manifested in the L1 grammar. Given that a language-particular grammar forms a subset of all possible grammars that are subject to universal computational constraints, 'direct UG-access' and 'UG-access via the L1' positions bear the same empirical signature (Dekydtspotter *et al.*, 1997).

In relation to remnant movement, however, the conceptual division of the Minimalist Program between a universal computational component, C_{HL}, and a language-particular functional lexicon leads to a well-defined perspective on the distinction between 'direct UG-access' and 'L1 mediation'. Let us consider the learning tasks for English learners and Japanese learners of German in this regard.

As argued above, the relevant typological differences between English, on the one hand, and German, on the other hand, reduce to the availability of a strong [scr] feature: English does not provide this feature in its functional lexicon, and hence provides no matching feature for a fronting operation different from A'-movements like topicalization and *wh*-movement. Yet, the universal computational principle Unambiguous Domination (UD) crucially depends on a differentiation of the type of features involved in remnant A'-movements. In order for English learners of German to show exactly the judgement pattern in (10), they must acquire knowledge of the properties of the [scr] feature, i.e., its movement type and specific checking sites. This information is neither encoded in the L1 English functional lexicon, nor does it follow from the invariant principles of the computational system. Hence,

'access to UG through the L1' would be insufficient for English learners of German, because the availability of UD alone does not yield the correct grammaticality contrasts.

By contrast, Japanese learners could in principle accommodate the pattern in (10) by mapping German syntax onto their isomorphic L1 featural inventory in this particular domain of grammar, i.e., [scr], [top], [wh].[5] In this case, the target-like grammaticality distinctions follow from the operation of UD. Thus, investigating whether the behaviour of English and Japanese learners of German is parallel on remnant movement can yield insight into the extent of UG involvement in L2 acquisition and the role of the native language.

IV Models of second language acquisition and their hypotheses

This study tests whether advanced English and Japanese learners know options and restrictions of word-order variation in German. For convergence, L2 learners need to know:

- that German allows scrambling;
- that German allows remnant movement; and
- that there are constraints on remnant movement.

Due to distinct L1 backgrounds, the learning tasks for the two groups differ: On the assumption that universal computational principles operate in IL grammars, English learners have to acquire a functional scrambling feature to arrive at target-like knowledge, whilst Japanese learners could have recourse to the L1 array of features in this area of grammar.

Of the approaches to L2 acquisition cited in the introduction, I consider the predictions of the so-called No Access, Partial Access and Full Access models. No Access models hold that general cognitive strategies guide L2 acquisition exclusively once UG has maturationally become unavailable. These strategies are predominantly, on the one hand, pattern accumulation on the basis of the frequency of individual construction types in the input and by analogy to the L1 grammar

[5]A learnability problem for Japanese learners of German arises for the delearning of scrambling across finite clause boundaries (long scrambling); for details and experimental results, see Hopp, 2002.

(Bley-Vroman, 1997) and, on the other hand, low-level analysis of linear surface strings in the TL input (Meisel, 1997). Section II demonstrated that neither of these strategies can ensure convergence on the L2 matrix of movement options. No Access models thus predict that L2 learners should fail to make robust discriminations along UG-governed lines, regardless of L1 background. Partial Access models maintain that the instantiation of grammatical properties in the L1 is a necessary prerequisite for successful acquisition in the L2. A recent version of the Partial Access approach claims that a specific subpart of UG, namely uninterpretable features in the functional lexicon, becomes immune to reparametrization after the offset of a critical period (Tsimpli and Roussou, 1991; Hawkins and Chan, 1997). Provided the functional array of the TL exceeds the L1 in the number of features or the number of feature values, learners are predicted never to converge on the morphosyntactic reflexes in the TL grammar. Rather, learners emulate the TL grammar by superficially (re)modelling their defective IL grammars on the input. Due to lack of input information of the ungrammaticality of remnant scrambling, the language groups are thus predicted to exhibit knowledge of UD effects with scrambling aligned with their L1s. Finally, Full Access models (Epstein *et al.*, 1996; Schwartz and Sprouse, 1996) hypothesize that learners can ultimately activate features and access computational principles irrespective of their realizations in the L1. Accordingly, both advanced English–German and Japanese–German IL grammars are expected to reflect TL properties.

V The study

1 Materials

The experiment investigated two paradigms of (remnant) movement in German, the infinitival paradigm as in (10) repeated below and an analogous DP paradigm as in (11).

10) Infinitival paradigm
    a. *Scrambling of complete phrase:*
    Ich glaube, dass [den Wagen zu reparieren]$_1$ Peter schon t$_1$ versucht hat.
    I think that the car to repair Peter already tried has
    'I think that Peter already tried to repair the car.'

b. *Topicalization of complete phrase:*
[Den Wagen zu reparieren]$_1$ hat Peter schon t_1 versucht.
The car to repair has Peter already tried

c. *Remnant topicalization across-scrambled phrase:*
[t_1 Zu reparieren]$_2$ hat Peter [den Wagen]$_1$ schon t_2 versucht.
To repair has Peter the car already tried

d. *Remnant topicalization (across-scrambled phrase) across finite clause boundary:*
[t_1 Zu reparieren]$_2$ glaube ich [t'_2 hat Peter [den Wagen]$_1$ schon t_2 versucht].
To repair think I has Peter the car already tried

e_1. *Remnant scrambling across short-scrambled phrase:*
* Ich glaube, dass [t_1 zu reparieren]$_2$ Peter [den Wagen]$_1$ schon t_2 versucht hat.
I think that to repair Peter the car already tried has

e_2. *Remnant scrambling across medium-scrambled phrase:*
* Ich glaube, dass [t_1 zu reparieren]$_2$ [den Wagen]$_1$ Peter schon t_2 versucht hat.
I think that to repair the car Peter already tried has

f. *Remnant topicalization across topicalized phrase:*
* [t_1 Zu reparieren]$_2$ glaube ich [den Wagen]$_1$ hat Peter schon t_2 versucht.
to repair think I the car has Peter already tried

The set of sentences in (11) charts options and restrictions for PP extraction from complex indefinite DPs. The movement types are homologous to the infinitival paradigm, except that there is no example of remnant scrambling across medium scrambling – as in ($10e_2$) – in the DP paradigm. This type would have the same surface order as intact scrambling (11a).

11) DP paradigm

a. *Scrambling of intact phrase:*
Ich denke, dass [einen Film über Frankreich]$_1$ Martin gestern t_1 gesehen hat.
I think that a film about France Martin yesterday watched has

b. *Topicalization of complete phrase:*
[Einen Film über Frankreich]$_1$ hat Martin gestern t_1 gesehen.
A film about France has Martin yesterday watched

c. *Remnant topicalization across scrambled phrase:*
[Einen Film t_1]$_2$ hat Martin [über Frankreich]$_1$ gestern t_2 gesehen.
A film has Martin about France yesterday watched

d. *Remnant topicalization (across scrambled phrase) across finite clause boundary:*
[Einen Film t_1]$_2$ denke ich hat Martin [über Frankreich]$_1$ gestern t_2 gesehen.
A film think I has Martin about France yesterday watched

e_1. *Remnant scrambling across short-scrambled phrase:*
* Ich denke, dass [einen Film t_1]$_2$ Martin [über Frankreich]$_1$ gestern t_2
I think that a film Martin about France yesterday
gesehen hat.
watched has

f. *Remnant topicalization across topicalized phrase:*
* [Einen Film t_1]$_2$ denke ich [über Frankreich]$_1$ hat Martin gestern t_2 gesehen.
A film think I about France has Martin yesterday watched

To test these contrasts, a 74-item grammaticality task was devised with three tokens for each sentence type in (10) and (11) and 35 filler items. Overall, 40 items were grammatical and 34 items ungrammatical.

A particular problem in eliciting judgements on optional constructions arises from their different degrees of acceptability, i.e., markedness. Optionality in word order in German is not manifested in binary terms of grammaticality, but gives rise to gradient grammaticality determined by a varied set of syntactic, phonological and pragmatic factors (Lenerz, 1977; Uszkoreit, 1986; Müller, 1999). Consequently, acceptability judgements in this regard are likely to be affected by preference rankings involving the conscious or subdoxastic comparison of alternative linearizations, so that rejection of a given linearization need not reflect its ungrammaticality *per se*, but rather its decreased acceptability compared to a less marked variant. Previous grammaticality judgement experiments on word-order optionality in Spanish and German (Liceras, 1993; Schreiber and Sprouse, 1998; Prévost, 1999) found that native and nonnative speakers display unstable judgements on items presented in isolation or items solely embedded in written contexts. Schreiber and Sprouse (1998) tested high-intermediate L1 English learners of German on a subset of the items in (10) and (11) in a written grammaticality judgement task that embedded items in contexts. They find low overall acceptance rates of grammatical sentences and indiscriminate judgements on grammatical and ungrammatical sentences in the subset of the DP paradigm they tested. It may be argued that these inconclusive findings are due to the format of the task, which insufficiently controlled for the discourse dependence of noncanonical word orders through intonation.

In the context of the present experiment, pilot studies with native speakers of German ($n = 74$) showed that the level and consistency in judgements of marked, discourse-dependent word orders rises considerably when the items are preceded by discourse contexts and explicit intonational information is provided. The bimodal presentation of items in discourse contexts systematically increased acceptance of both grammatical and ungrammatical items (for details, see Hopp, 2002). In addition to facilitating the acceptance of marked noncanonical orders, providing intonation made ill-formed sentences seem more acceptable, too. Importantly, this overall rise in acceptance shows that

the intonation chosen in the bimodal condition was appropriate, and that intonation did not provide clues to discriminate between grammatical and ungrammatical items. Nevertheless, a certain degree of variability in judgements remained in the bimodal condition, which suggests that context and intonation are ultimately insufficient to neutralize the gradient grammaticality of noncanonical word orders in German.

For the main experiment, each item was embedded in an appropriate discourse context and presented bimodally in order to minimize effects of gradient grammaticality. There were no scrambled orders in the contexts to avoid syntactic priming effects. All content words of the test items occurred in the preceding contexts to eliminate potential difficulties with novel words. To avoid potential processing difficulties arising from syntactic ambiguity, all DPs bore unambiguous case inflection. Lexical items and sentence length were held constant in every paradigm. Thus, we could compare minimal septuples (infinitival paradigm) and sextuples (DP paradigm) of lexically identical sentences which only differed with respect to word order. The items were presented in two distinct pseudo-randomized orders that spread similar contexts and constructions as far apart as possible. As for intonation, all contexts and items were read at a moderate pace by a male speaker of standard German (High German). They were digitally recorded and edited before being stored on audio CDs. In both paradigms, sentences of the same lexical items were matched for reading time (±100 ms). The test items were presented twice in identical form, divided by a three-second pause. After the second presentation of the test item, subjects were given 11 seconds to make a judgement before the CD automatically moved on to the next track.

Extensive instructions and practice items preceded the task. Responses were recorded on a five-point rating scale ranging from '–2' ('not possible') to '+2' ('possible'); '×' for 'don't know'. Representative examples of test items from the infinitival and the DP paradigm are given in (12) and (13).

12) Infinitival paradigm:
Andreas and Annette are talking about Peter's car, which broke down recently. It looks as though Peter was going to take his car to the garage. Annette knows that Peter almost always repairs his car himself. That is why she asks Andreas whether

Peter wants to repair his car himself this time as well. Andreas thinks that Peter already tried to do that, but he is not certain. He replies:

Ich glaube, dass den Wagen zu reparieren Peter schon versucht hat.
'I think that Peter already tried to repair the car.'

Is this sentence possible in this context? −2 −1 +1 +2 ×

13) DP paradigm:
Claire and Jutta talk about Martin. Jutta knows that Martin is a great fan of all things French. He often listens to French songs, reads books about France and enjoys watching films about France. Therefore, Jutta asks what Martin did with respect to France yesterday. Claire replies:

Einen Film hat Martin über Frankreich gestern gesehen.
'Martin saw a film about France yesterday.'

Is this sentence possible in this context? −2 −1 +1 +2 ×

2 *Subjects*

Twenty-six English-speaking and 13 Japanese-speaking learners of German took part in the experiment, as well as 26 native speaker controls. All of the native German and nonnative subjects were resident in Germany at the time of testing with the exception of sixteen L1 English speakers who were final-year students of German at British universities. The mean age of the L2 subjects was 34.9 years. The nonnative subjects had been exposed to German for at least four years, although most subjects had had considerably longer exposure to German (overall mean: 15.9 years). Each nonnative subject had spent at least six months in a German-speaking country, and most subjects had lived in a German-speaking context for several years (overall mean: 6.7 years). Table 1 presents detailed information about the nonnative speakers' backgrounds.

The 39 nonnatives were allocated to three proficiency groups on the basis of a 40-item cloze test: high intermediate (scores from 14 to 25), advanced (scores from 26 to 35) and very advanced (scores above 36).[6] *T*-tests indicate significant differences between levels within each language group ($p < 0.01$). Comparison across language groups does not yield significant differences between respective levels.

[6]Unfortunately, there were not enough subjects for a very advanced Japanese group. Note also that the grouping leads to differences in the mean length of exposure for the matched groups, E_1 and J_1. Subjects in E_1 have had considerably longer exposure than subjects in J_1. This discrepancy is counterbalanced by the fact that most subjects in E_1 have a considerably lower length of residence in a German-speaking country than their matched Japanese counterparts. A similar reasoning applies for groups E_2 and J_2; for details, see Hopp, 2002.

Table 1 Subject information by group (arithmetic mean is given in parentheses)

	Native group (n = 26)	E_1 (n = 7) (English high intermediate)	E_2 (n = 13) (English advanced)	E_3 (n = 6) (English very advanced)	J_1 (n = 8) (Japanese high intermediate)	J_2 (n = 5) (Japanese advanced)
Cloze test (max. 40)	35–40 (37.2)	14–25 (19.3)	26–33 (28.7)	36–40 (39.2)	14–24 (18.8)	27–34 (30.4)
Length of exposure (years)		9.5–20 (12.5)	9.5–40 (16.7)	16–45 (29.2)	4–9 (6.4)	16–35 (22)
Length of residence (years)		0.8–18 (3.3)	0.5–31 (6.2)	4–33 (15.2)	1–9 (4.8)	8–25 (17.2)
Age of onset		11–35 (15.6)	11–14 (12.5)	13–14 (13.3)	11–39 (24.6)	13–40 (23.0)
Age		22–68 (29.7)	22–53 (27.8)	27–60 (15.7)	17–43 (31.9)	37–56 (45.4)

3 Results

The principal findings of the study can be summarized as follows:

- Irrespective of proficiency level or L1 background, all nonnative groups make robust relative distinctions between licit (remnant) movement and illicit remnant scrambling.
- Each L2 speaker makes reliable individual distinctions between licit (remnant) movement and illicit remnant scrambling.
- The Japanese groups show higher levels of acceptance on scrambling than the respective English groups ($p < 0.05$), whereas the Japanese and English groups' performance on topicalization is statistically indistinguishable ($p > 0.05$).

In all groups, Wilcoxon Signed Rank tests for paired samples yield robust relative contrasts between grammatical and ungrammatical types of (remnant) movement. Although comparisons were computed on the whole dataset, I report only the percentages of positive judgements. Given the very low amount of 'not sure' judgements (0.3% for all groups), the rejection rates are virtually linearly dependent on the acceptance rates. In the following, I focus on the relative contrasts in acceptability judgements between grammatical and ungrammatical items. I do not discuss in detail within-group differences in the absolute level of judgements that are caused by the graded grammaticality of noncanonical word orders.

Table 2 charts the responses of the English and native groups in the infinitival paradigm (10). In the infinitival paradigm, a sharp contrast arises between the acceptance rates of illicit remnant scrambling ($10e_1$ and $10e_2$) and licit scrambling of complete phrases (10a) for all nonnative groups at a significance level of $p < 0.02$ or less. Similarly, the two types of illicit remnant scrambling (10e) are accepted at significantly lower levels than the licit types of remnant topicalization (10c and 10d). Finally, illicit remnant topicalization (10f) is tolerated considerably less frequently compared to its licit counterparts (10c and 10d).

Considering the L1 English proficiency groups separately, we find a relatively high degree of conformity. Intergroup Mann–Whitney comparisons within the L1 English groups yield statistically significant differences solely for licit remnant movement: Among the three English

Table 2 Acceptance of sentence types in the infinitival paradigm (10) for English and German groups (judgements of [+1] and [+2] grouped together; percentages in parentheses)

Infinitival paradigm	E_1 (n = 7) (high intermediate English group)	E_2 (n = 13) (advanced English group)	E_3 (n = 6) (very advanced English group)	Natives (n = 26)
10a) Scrambling	13/21 (61.9)	25/39 (64.1)	12/18 (66.7)	62/77 (80.5)
10b) Topicalization	19/21 (90.5)	35/39 (97.4)	17/18 (94.4)	73/78 (93.6)
10c) Remnant topicalization (scr)	8/14 (57.1)	10/26 (38.5)	3/11 (27.3)	31/52 (59.6)
10d) Long remnant topicalization (scr)	10/14 (71.4)	10/26 (38.5)	4/11 (36.4)	22/52 (42.3)
$10e_1$) *Remnant scrambling (s-scr)	1/21 (4.8)	4/39 (10.3)	2/17 (11.8)	10/78 (12.8)
$10e_2$) *Remnant scrambling (m-scr)	5/21 (23.8)	3/39 (7.7)	1/18 (5.6)	19/78 (24.4)
10f) *Remnant topicalization (top)	2/21 (9.5)	4/39 (10.3)	3/18 (16.7)	2/78 (2.6)

groups, the least advanced group, E_1, shows significantly divergent responses on long remnant topicalization in the infinitival paradigm to E_2 ((10d): $p = 0.042$) and E_3 ((10d): $p = 0.028$).[7] In view of the high degrees of similarity across proficiency and the small sizes of the individual non-native groups, I report all results from comparisons between nonnatives and natives for subjects collapsed by language group. In the infinitival paradigm, the performance of the English group does not vary statistically significantly from the native speakers on any sentence type ($p > 0.05$). The sole exception is the English groups' response on illicit remnant scrambling ($10e_2$), which English learners as a group accept at significantly lower levels than native controls.

The judgements by both Japanese groups exhibit patterns highly similar to the English groups and the German controls (Table 3): Judgements on licit scrambling (10a) differ significantly from illicit remnant scrambling ($10e_1$ and $10e_2$) for J_1 ($p < 0.001$) as well as J_2 ($p = 0.017$). Highly significant differences arise between topicalization (10b) and illicit remnant topicalization (10f) for both groups

[7]I submit that this construction-particular difference between proficiency levels is due to prescriptive bias, stylistic preferences or a conservative judgement strategy to avoid making errors adopted by the more advanced groups (see also Coppieters, 1987). The important observation remains that all groups draw a strong relative distinction between licit and illicit movement types.

Table 3 Acceptance of sentence types in the infinitival paradigm (10) for Japanese and German groups (judgements of [+1] and [+2] collapsed grouped together; percentages in parentheses)

Infinitival paradigm	J_1 (n = 8) (high intermediate Japanese group)	J_2 (n = 5) (advanced Japanese group)	Natives (n = 26)
10a) Scrambling	22/24 (91.7)	11/15 (73.3)	62/77 (80.5)
10b) Topicalization	23/24 (95.8)	13/15 (86.7)	73/78 (93.6)
10c) Remnant topicalization (scr)	7/16 (43.8)	5/10 (50)	31/52 (59.6)
10d) Long remnant topicalization (scr)	6/16 (37.5)	5/10 (50)	22/52 (42.3)
$10e_1$) *Remnant scrambling (s-scr)	0/24 (0)	1/15 (6.7)	10/78 (12.8)
$10e_2$) *Remnant scrambling (m-scr)	4/24 (16.7)	4/15 (26.7)	19/78 (24.4)
10f) *Remnant topicalization (top)	2/24 (8.3)	2/15 (13.3)	2/78 (2.6)

($p < 0.001$). The high intermediate group J_1 also displays significant differences ($p < 0.05$) in responses between the two types of licit remnant movement and the three types of illicit remnant movement. For the comparatively small advanced group J_2, some of these differences fail to reach statistical significance despite robust aggregate distinctions. Intergroup Mann–Whitney comparisons do not measure any significant differences on any sentence type between the two Japanese groups or between the Japanese groups and the German group.

In the DP paradigm (Tables 4 and 5), the significant contrasts between licit and illicit sentence types are replicated even more prominently. In fact, nonnatives tend to establish stronger oppositions between grammatical and ungrammatical sentence types than native speakers. In the responses of the three English groups, illicit remnant scrambling (11e) is found to differ highly significantly ($p < 0.001$) from all other licit types of movement. Intergroup Mann–Whitney comparisons between the three English groups fail to uncover significant differences on any sentence type. Intergroup comparisons between the English groups and the German group yield significant differences for some grammatical sentence types, namely, scrambling ($p = 0.003$), remnant ($p = 0.003$) and long remnant topicalization ($p = 0.003$). In each case, however, the aggregate English group exhibits significantly higher acceptance of these grammatical types than the native control group. I discuss the

Table 4 Acceptance of sentence types in the DP paradigm (11) for English and German groups (judgements of [+1] and [+2] grouped together; percentages in parentheses)

DP paradigm	E_1 (n = 7) (high intermediate English group)	E_2 (n = 13) (advanced English group)	E_3 (n = 6) (very advanced English group)	Natives (n = 26)
11a) Scrambling	15/21 (71.4)	23/39 (59.0)	11/18 (61.1)	34/78 (43.6)
11b) Topicalization	21/21 (100)	38/39 (97.4)	18/18 (100)	77/78 (98.7)
11c) Remnant topicalization (scr)	17/21 (81.0)	34/39 (87.2)	13/17 (76.5)	48/77 (62.3)
11d) Long remnant topicalization (scr)	17/21 (81.0)	34/39 (87.2)	16/18 (88.9)	59/78 (75.6)
11e) *Remnant scrambling (s-scr)	4/21 (19.0)	4/39 (10.3)	0/17 (0)	4/78 (5.1)
11f) *Remnant topicalization (top)	10/21 (47.6)	23/39 (56.4)	10/18 (55.6)	45/78 (57.7)

Table 5 Acceptance of sentence types in the DP paradigm (11) for Japanese and German groups (judgements of [+1] and [+2] grouped together; percentages in parentheses)

DP paradigm	J_1 (n = 8) (high intermediate Japanese group)	J_2 (n = 5) (advanced Japanese group)	Natives (n = 26)
11a) Scrambling	22/24 (91.7)	12/15 (80.0)	34/78 (43.6)
11b) Topicalization	24/24 (100)	15/15 (100)	77/78 (98.7)
11c) Remnant topicalization (scr)	16/23 (69.6)	15/15 (100)	48/77 (62.3)
11d) Long remnant topicalization (scr)	14/24 (58.3)	11/15 (73.3)	59/78 (75.6)
11e) *Remnant scrambling (s-scr)	4/24 (16.7)	2/15 (13.3)	4/78 (5.1)
11f) *Remnant topicalization (top)	12/24 (50.0)	11/15 (73.3)	45/78 (57.7)

comparatively low acceptance rates by the native speakers in Section IV, subsection 2.

For the two Japanese groups (Table 5), ungrammatical scrambling in the DP paradigm is statistically highly different from the licit types of (remnant) movement ($p < 0.002$). Statistically significant differences between groups can be measured only with respect to remnant topicalization in the DP paradigm ((11c): $p = 0.020$), which the high intermediate group J_1 accept to a lower overall degree compared to the advanced group J_2. In comparison with native speakers of German, L1

Japanese learners accept scrambling of indefinite DPs at significantly higher levels than native speakers ($p < 0.001$).

Like the native controls, nonnatives equally accept the supposedly ungrammatical remnant topicalization pattern (11f). I submit that some of the native and nonnative speakers analyse the construction (11f) as intact topicalization involving a parenthetical finite clause (e.g., *denke ich* 'I think'). This would seem possible in a pragmatic context where a speaker signals uncertainty about the referent of the PP by hedging the topicalized constituent in mid-sentence. Reis (1995) notes that such parenthetical insertion is (pragmatically) main-clause bound and may have to be relegated to discourse grammar. Arguably, an analysis along these lines receives support in view of the lower acceptance of this sentence type in the exclusively written pilot version of the task (Hopp, 2002).[8]

To recapitulate, the group results bear out that grammatical and ungrammatical sentence types are consistently distinguished in the two syntactic paradigms despite the superficial dissimilarity of these paradigms. All five nonnative groups manifest robust and native-like relative contrasts between grammatical and ungrammatical sentence types. Where behaviour diverges, nonnative judgements are not unilaterally deficient compared to the control group. Moreover, L2 learners manifest native-like relative distinctions between sentence types irrespective of L1, and the knowledge of the grammaticality contrasts in (10) and (11) is independent of the proficiency levels tested. Finally, all groups display low absolute levels of acceptance of ungrammatical items in spite of positively biased contextual and intonational presentation.

4 Variation across individuals

In view of the variability in acceptability of noncanonical word orders, it is essential to consider individual judgement patterns. Examining the relative distinctions between sentence types for each subject allows us to observe the degree of inter-subject variation in judgements and hence delineate the range of the hypothesis space explored by individual learners. If the relative contrasts attested at the aggregate level are

[8]An anonymous reviewer points out that one could examine whether subjects adopt a parenthetical interpretation of *I think* by using nonfactive verbs such as *doubt*, which do not allow for parenthetical insertion.

robust across individuals, they should translate into clear relative discriminations made by each subject.

I first consider the relative discrimination between licit of complete phrases scrambling and illicit remnant scrambling; I then turn to the distinction between licit remnant movement and illicit remnant movement types. Figure 1 shows the distribution of native subjects as regards their relative acceptance of licit scrambling compared to their acceptance of illicit remnant scrambling in the infinitival and DP paradigm grouped together. For instance, the value of 20% on the horizontal axis means that scrambling is accepted at a rate 20% higher than remnant scrambling. The vertical axis denotes number of subjects.

Although the spread of the distribution in Figure 1 indicates that the graded grammaticality of scrambling may result in varying degrees of acceptance by native speakers, scrambling of complete phrases is overwhelmingly judged as more acceptable than illicit remnant scrambling. Only two subjects in the native sample do not prefer licit scrambling over illicit (remnant) scrambling or treat the two types on a par. By all other 24 subjects, illicit remnant scrambling is rated as worse than licit scrambling, and 77% of natives exhibit a reliable distinction at a rate of at least 30%, with the overall group mean at 45%.[9]

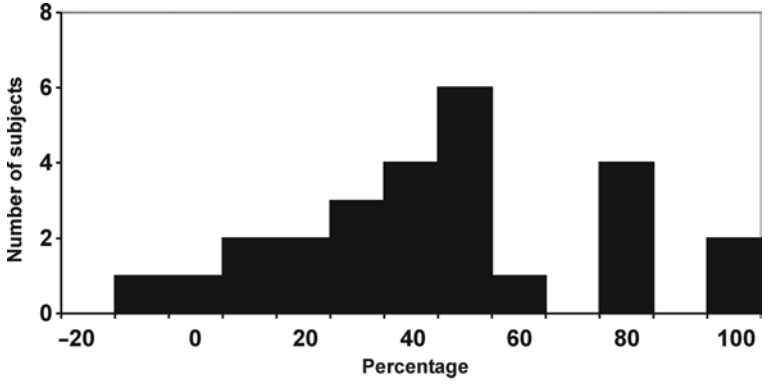


Figure 1 Native group (n = 26): relative discrimination of licit scrambling of complete phrases (10a and 11a) vs. illicit remnant scrambling ($10e_{1,2}$ and 11e) across paradigms

[9]I follow Dekydtspotter *et al.* (1997) in arbitrarily adopting a 30% criterion as denoting a reliable distinction rate.

For expository convenience, all nonnative groups are shown in Figure 2 grouped together. The skewed pattern for the nonnative groups in Figure 2 illustrates that all subjects draw a unilateral distinction between grammatical and ungrammatical scrambling constructions.[10] The fact that 92% exhibit a contrast of at least 30% between the types (overall mean: 54%) shows that the distinction is robust across subjects. Overall, the skewed distribution in Figure 2 evidences that L2 behaviour is highly systematic in relation to scrambling in German, since no single learner accepts ungrammatical remnant scrambling over grammatical scrambling.

Turning to the different types of remnant movement, Figure 3 illustrates for the native speakers that all subjects distinguish robustly between licit remnant movements and illicit remnant movements, i.e., remnant scrambling and remnant topicalization across topicalization: 84% do so at a rate of minimally 30%, and the group mean is 49%. Figure 4 displays a comparable picture for the L1 English and Japanese sample populations grouped together. All subjects fall into the positive range, i.e., they all prefer grammatical over ungrammatical instances of remnant movements. 89.7% of subjects make a reliable distinction of at

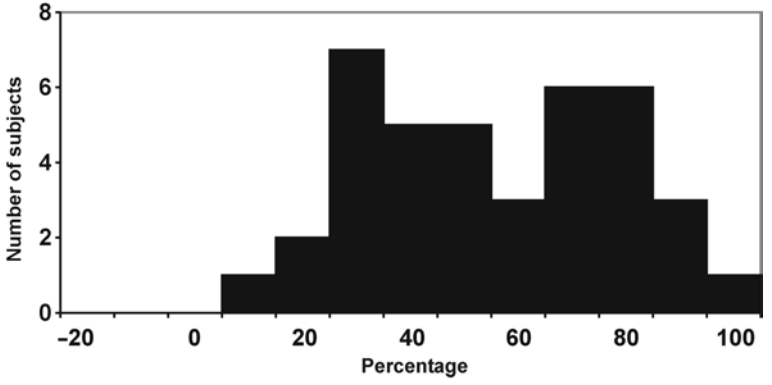


Figure 2 Nonnative groups (n = 39): relative discrimination of licit scrambling of complete phrases (10a and 11a) vs. illicit remnant scrambling ($10e_{1,2}$ and 11e) across paradigms

[10]This also holds if one computes relative differences for the infinitival and the DP paradigm separately.

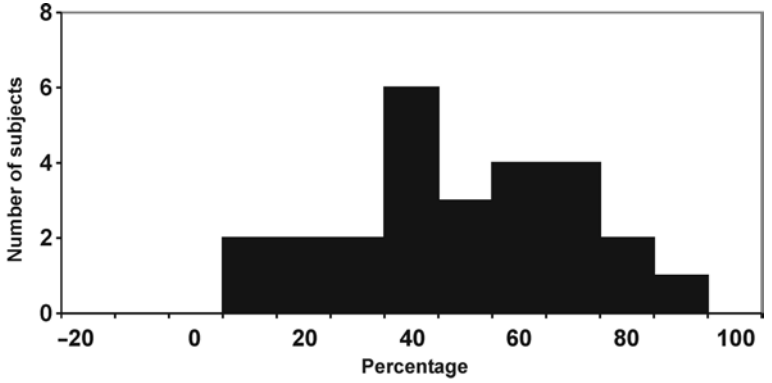


Figure 3 Native group (n = 26): relative discrimination of licit remnant topicalization (10c–d and 11c–d) vs. illicit remnant movements (10e–f and 11e–f) across paradigms

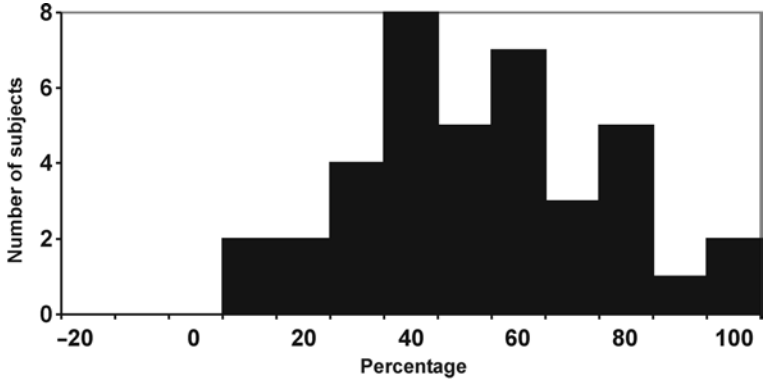


Figure 4 Nonnative groups (n = 39): relative discrimination of licit remnant topicalization (10c–d and 11c–d) vs. illicit remnant movement (10e–f and 11e–f) across paradigms

least 30%, with an overall mean distinction of 53% between licit and illicit remnant movements.

To summarize, examining the individual response patterns allows us to establish a more robust distinction between ungrammatical constructions and marked grammatical constructions. If the difference between these two types were merely a statistical artifact in the aggregate data, one would expect greater inter-subject variation than observable in Figures 1 to 4. In particular, one would expect some evidence of reversed preference patterns between licit and illicit items if individual

learners adopted idiosyncratic response strategies. Yet, for every nonnative subject, illicit types of remnant movement are treated differently from licit cases of remnant movement or scrambling in a unidirectional way. Although inter-subject variation exists as to the degree of differentiation, not a single subject treats licit and illicit types equally or accepts illicit sentence types at higher levels than their licit counterparts. Importantly, these distinctions are not construction specific, as robust acceptability contrasts are made for both scrambling and remnant movement constructions. At least for the English groups, the IL behaviour thus requires grammatical IL representations distinct from the grammatical configuration responsible for the L1 pattern. Furthermore, the spectrum of judgements across individuals is highly and systematically constrained. Such absence of variation is unexpected in general learning scenarios, as individuals' hypotheses should not *a priori* be skewed in a typical population sample. Rather, these findings indicate that the hypothesis space explored in the L2 acquisition of German word order options by learners of typologically different L1s is narrowly restricted by the constraints operative in child L1 acquisition.

5 L1 differences

Whereas no significant differences are attested between the L2 groups for knowledge of UG-specified ungrammaticality within the grammatical sentences, there is a systematic and construction-specific asymmetry between the two nonnative groups. We can measure a robust interaction effect between L1 and aggregate responses in relation to scrambling. Comparing the two licit types of intact category movement, scrambling (10a and 11a) and topicalization (10b and 11b), we note that the English groups accept scrambling of complete phrases at significantly lower rates than the respective Japanese groups ($p < 0.05$) in both paradigms (Figures 5 and 6). No asymmetry obtains for topicalization of complete phrases in either paradigm ($p > 0.05$).

As shown in Figure 5, the differences between the higher intermediate groups, E_1 and J_1, attain statistical significance for both paradigms ($p < 0.05$). Although the relative asymmetries persist at higher levels of proficiency (see Figure 6), the absolute differences between L1 groups on scrambling diminish, pointing to gradual convergence. In comparison

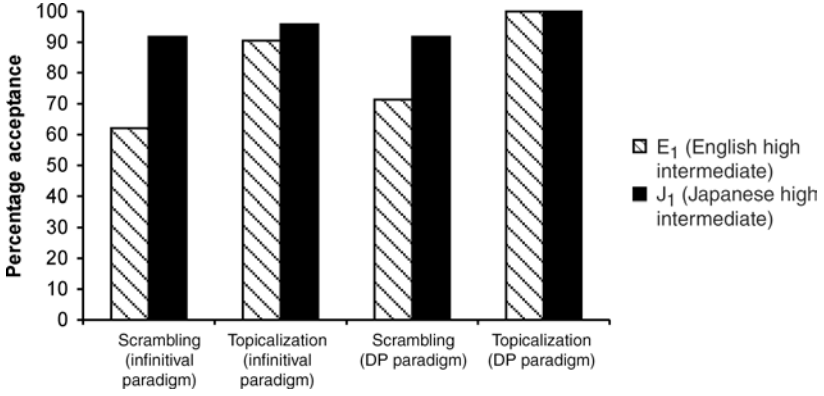


Figure 5 Comparison of aggregate acceptance of scrambling and topicalization of complete phrases by L1 groups per paradigm (for high intermediate groups)

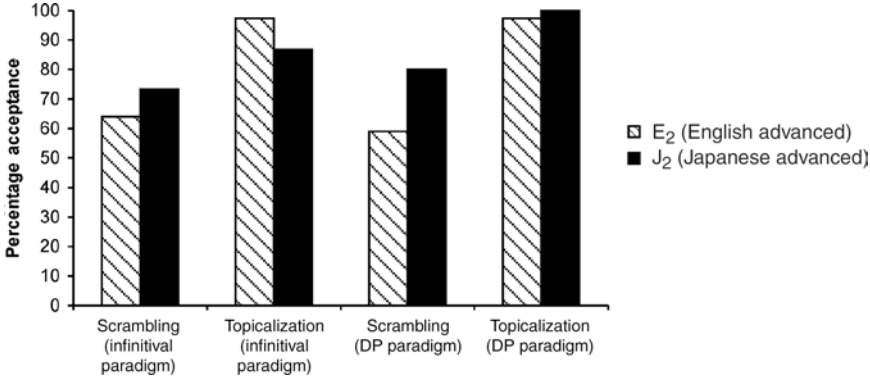


Figure 6 Comparison of aggregate acceptance of scrambling and topicalization of complete phrases by L1 groups per paradigm (for advanced groups)

to J_2, scrambling is accepted by E_2 at markedly lower rates, particularly in the DP paradigm, in which the 80% acceptance rate for J_2 contrasts with rates around 60% for E_2.

By contrast, there is no interaction between L1 and level of acceptance on any of the four grammatical types of topicalization for any of the proficiency groups ($p > 0.05$). A significant difference does also arise for L1 groups grouped together across proficiency. English subjects as a group accept scrambling at significantly lower rates ($p < 0.05$) than

all Japanese subjects as a group in the infinitival paradigm, and the difference between L1 groups is highly significant in the DP paradigm ($p < 0.001$). Statistical analysis does not yield significant differences for any other sentence type ($p > 0.05$).

Given that scrambling out of coherent infinitives and indefinite DPs occurs very rarely in the input, it is highly unlikely that differences in input frequency lead to the differences in acceptance rates between the English and Japanese groups. We can equally discount general stylistic preferences or judgement strategies as an explanation, because subjects in E_1 accept all other sentence types at higher or similar absolute rates compared to the respective Japanese group, J_1. Consequently, the intergroup differences in responses to scrambling constructions are dependent on the variable of L1 background. In spite of the advanced proficiency levels of the subjects, the observed interaction effect between acceptance of scrambling and L1 can thus be logically related to typologically distinct initial states of IL development (Schwartz and Sprouse, 1996: 67).

VI Discussion

1 UG vs. judicious strategies

The central questions in this study are whether adult L2 learners converge on TL properties despite Poverty of the Stimulus, and whether convergence depends on analogous properties in the L1 grammar. The results from a grammaticality judgement task show that word-order variation in IL German does not manifest options outside those permitted by UG. All English and Japanese learners robustly replicate the pattern of relative judgements by native German speakers.

Empirical observation of highly regimented L2 behaviour requires explanation, in particular the fact that relative discriminations made by groups of typologically different L1s are very similar, although they face distinct learnability tasks. Learning mechanisms that centre around statistical regularities in the input and surface analogies with the L1 (Bley-Vroman, 1997; Meisel, 1997) fail to produce the relevant discriminations. It is highly unlikely that:

- the small differences between the minimal frequency of the grammatical types and the non-occurrence of ungrammatical sentences

enable L2 learners to draw robust and inter-individually systematic distinctions along UG-governed lines and, consequently, that:

- a systematically circumscribed matrix of licit optional word orders – as in (10) and (11) – should form a stable part of nonnative language systems.

Rather, the present results clearly point to the involvement of grammatical knowledge going beyond input properties. As often argued, L2 learners could attempt to compensate for the lack of a UG-constrained hypothesis space by patterning their behaviour on their L1s. Construed extensionally, such an assumption is demonstrably incorrect, since the drastic surface dissimilarities between English and Japanese would give rise to markedly dissonant hypotheses for the English and Japanese groups. Furthermore, we would expect a substantial degree of inter-subject variation depending on whether and how individual learners draw analogies or match patterns with their L1 (Bley-Vroman, 1990). Finally, one would expect to see a proficiency effect as regards accuracy of judgements even at advanced levels, since greater knowledge of the TL should lead more advanced learners to refer to L1 properties less frequently than less advanced learners. Disconfirming these predictions, the uniform behaviour across syntactic paradigms, subjects, L1s and proficiency groups investigated testifies that neither low-level extrapolation from (the frequency of) structural patterns (Bley-Vroman, 1990), nor linear surface sequences in the input (Meisel, 1997), nor L1 surface analogy guides L2 judgements.

Construed intensionally, the argument from L1 analogy would take the line that English grammar does not lend itself to a ready accommodation of German movement options, whereas Japanese word-order freedom offers higher chances of successfully restricting optionality. Therefore, in principle, recourse to UG beyond the L1 instantiation might not be necessary for Japanese learners, and it might not be possible for English learners. Yet, evidence of this kind of asymmetry is conspicuously absent in the data. Both groups evidence that UG-specified knowledge is available independently of the grammaticality status of ostensibly equivalent constructions in the L1. Given Minimalist design, moreover, restricting IL grammatical knowledge to L1 aspects independent of UG becomes a narrowly delimited option.

The discriminations between movement types in German have been shown to follow from computational principles that impose universal constraints on overt constituent displacement. These interact with language-specific functional features in licensing a subset of word-order options (Section II). Carving out fragments of L1 grammatical knowledge from UG in order to account for L2 behaviour (Clahsen and Muysken, 1989; Schachter, 1996) thus becomes a highly suspect move, since the L1 grammar fully instantiates UG principles. According to Minimalist architecture of the language faculty, any intensional treatment of L2 grammars necessarily entails the involvement of a substantial subset of UG properties.

In L1 acquisition, these domain-particular configurations ensure convergence on TL-properties in spite of scant evidence of word-order restrictions in the linguistic input. Since the same learnability task obtains in the L2 acquisition of word-order constraints not instantiated in the L1, the finding that advanced nonnative L1 English speakers perform like native speakers strongly points to the involvement of identical domain-specific mechanisms in native and nonnative acquisition.

Fashioning a Partial Access model within the confines of UG, Hawkins and Chan (1997) argue that uninterpretable features not present in the L1 become inaccessible to postpubescent learners. Rather, it is argued that adult L2 learners categorize lexical and morphophonological items of the TL within the functional repertoire of their L1. Since there is no scrambling feature in their functional lexicon, Hawkins and Chan (1997) would predict that English learners (mis)categorize German scrambling as topicalization. Topicalization is the closest analogue to German scrambling in English, because it equally leads to the optional fronting of objects. On such an analysis, English learners could accommodate surface word-order variation within their defective IL grammars. Whilst generally successful in emulating L2 patterns, this strategy would break down for remnant movement. Were scrambling to be classified as topicalization, the licit types of remnant movement, where scrambling feeds topicalization, should be disallowed by the principle of Unambiguous Domination as illicit instances of topicalization across topicalization. In fact, all types of remnant movement in (10) and (11) should be treated on a par by English learners, since the relevant featural distinction is allegedly not

available in their lexicon. This prediction is clearly incompatible with the findings. The aggregate results demonstrate a clear asymmetry in the English IL judgements. For proficiency groups and syntactic paradigms grouped together, the mean acceptance by English learners of licit remnant topicalization (10c, d and 11c, d) is 68.5%. By contrast, mean acceptance of illicit remnant scrambling (10 and 11e) is 10.3%, and mean acceptance of illicit remnant topicalization in the infinitival paradigm (10f) reaches 11.5% (see Section V). These statistically significant differences between acceptance of licit and illicit remnant movement types document that English IL grammars do in fact encode a distinction between scrambling and topicalization. In short, they are treated as distinct movement types. In this study, there is thus no evidence that access to the functional UG lexicon of features is impaired in adult L2 acquisition.

2 L1 effects

The present results suggest that native-like relative discriminations can be attained, regardless of whether the grammatical phenomena are instantiated in the native language. Having thus established the necessity for the full involvement of UG in the behaviour of the L2 groups, let us turn to differences in the groups' performances. The L1 groups differ systematically at the level of absolute acceptance of licit scrambling (Figures 5 and 6). Namely, the high intermediate English learners of German (E_1) demonstrate significantly lower approval of scrambling in both paradigms (10 and 11a) than the corresponding Japanese group (J_1). At higher levels of proficiency, the behaviour of English and Japanese learners becomes more similar, although a sizeable asymmetry remains in the DP paradigm. For topicalization, there are no significant differences between any of the groups. Under a No Transfer/Full Access account, which considers the L1 not to systematically affect IL representations at any point (e.g., Epstein *et al.*, 1996), groups of comparable proficiencies are expected to pattern alike, provided input is constant. Since it can be ruled out that frequency differences in the input are exactly parallel to L1 background, the observed interaction between acceptance of intact scrambling and L1 can be logically related to typologically distinct initial states of IL development in spite of advanced

proficiency. Unlike the Japanese learners, English learners have to restructure their IL grammars by allowing for scrambling across the subject.[11] In view of the low discourse frequency of scrambling, this restructuring seems to be subject to difficulty for the English groups, as suggested by the lower acceptance of this type of scrambling *vis à vis* the Japanese groups and also the native controls. This difficulty would seem to reflect the scarcity of scrambling in the TL input. Due to the infrequency of scrambling, Hopp (2004) argues that L2 learners have protracted problems identifying the target-like mapping between discourse function and the syntax of scrambling. It is important to note that the lower acceptance rates by the English groups reflect a quantitative, rather than a qualitative, difference. As argued above, were English learners not capable of restructuring their grammars to incorporate a scrambling feature at all, they would be unable to make the relevant discriminations in the remnant movement case, contrary to fact.[12]

In striking contrast to the lower judgement levels by English speakers, the Japanese groups' acceptance of intact scrambling significantly exceeds the acceptance levels by native speakers in the DP paradigm (11a). Aggregate acceptance of intact scrambling by J_1 is 91.7%, and 80% by J_2, whilst it patterns as low as 43.6% among the native controls (Table 5). This discrepancy becomes explicable from the extrasyntactic licensing conditions on scrambling of indefinites. In German, scrambling of DPs as in (14) is subject to a definiteness constraint, which renders scrambling of indefinites marginal (Abraham, 1986).

[11]In the technical terms of the Minimalist analysis in Section II, this means that English learners restructure their IL grammars by allowing for adjunction to AgrSP for the checking of a [scr] feature.

[12]An anonymous reviewer asks how this experiment actually tests whether L2 learners construct IL grammars constrained by UG given that discourse factors modulate scrambling. S/he goes on to say that 'it could be argued [that] acquiring these discourse properties/constraints form [*sic*] the core of the acquisitional task for the L2 learner (and the L1 child).' While it is true that acquiring discourse constraints constitutes an important part of acquiring (second) languages, discourse constraints cannot explain consistent behaviour in distinguishing between grammatical and ungrammatical sentences. Discourse constraints specify the appropriate use of grammatical constructions, i.e., when it is more or less felicitous to use a grammatical construction in linguistic context. They do not explain the grammatical regularities of syntax, which hold irrespective of discourse contexts. Indeed, pilot studies for this study identified the most appropriate discourse contexts for marked and ungrammatical scrambling sentences (Hopp, 2002). Yet, even in these felicitous discourse contexts, all native and nonnative groups robustly reject ungrammatical sentences at above 75%. These high levels of rejection and consistency clearly implicate discourse-independent grammatical knowledge.

14) ? Ich denke, dass [einen Film über Frankreich]$_1$ Martin gestern t_1 gesehen hat.
I think that a film about France Martin yesterday watched has
'I think that Martin watched a film about France yesterday.'

Indeed, the low overall acceptance of this type by native speakers (43.6%) bears out the operation of the definiteness constraint.

Neither Japanese group reflects this marginality in their judgements. Japanese does not encode definiteness grammatically (Saito, 1985), and scrambling is semantically vacuous (Saito, 1989). Hence, the Japanese learners' insensitivity to the definiteness effect appears to reflect transfer of L1 properties. These deviant L1 properties persist to advanced level due to the absence of restructuring information in the TL input. Compared to the rare global occurrence of scrambled definites in German, the virtual non-occurrence of scrambled indefinites is unlikely to provide sufficiently robust evidence for the Japanese learner to induce the definiteness restriction on German scrambling. Similarly to the English groups' performance on scrambling, the Japanese groups' divergence from the native speakers in this respect appears to be reflective of L1 transfer and ambiguity of the relevant input properties, rather than being a consequence of grammatical deficits, which categorically preempt convergence on the L2 (Hopp, 2004).

To summarize, although the findings suggest that the acquisition of 'new' formal features triggering movement is fully possible in IL grammars, whether a property is instantiated in the L1 can be demonstrated to be the source of quantitative differences in the performance, even of advanced L2 speakers. Thus, rather than lead to qualitative failure in the acquisition of German word-order optionality, as predicted by Partial Access models, L1 properties were shown to incur relative differences in the gradual restructuring to TL properties. These results support the Full Transfer/Full Access model (Schwartz and Sprouse, 1996), according to which L2 acquisition proceeds by the gradual and unimpaired restructuring from the L1 grammar mediated by UG constraints and triggered by positive TL input.

VII Conclusions

This study finds that nonnative speakers of distinct L1s demonstrate narrowly constrained native-like reflexes of UG-governed knowledge

despite Poverty of the Stimulus. Such evidence suggests that it is highly unlikely that maturational constraints enforce a radical asymmetry in knowledge states between L1 and adult L2 grammars. On Minimalist assumptions, the highly regimented L2 behaviour elicited in the present study across L1 groups points to the unimpaired availability of the computational principles and the parametric features of Universal Grammar in adult L2 acquisition. Notwithstanding qualitative parallelism in judgements between groups, this study found quantitative L1-specific divergences for scrambling. These differences in behaviour suggest that advanced learners have protracted difficulties identifying the semantic and information-structural correlates of syntactic reordering in the L2. Whether this is merely the result of impoverished input or whether this points to functional deficits at the interfaces of syntax with semantics and information structure, where discourse information is mapped onto word order, is for further research to elucidate (see Hopp, in progress).

Acknowledgements

I am greatly indebted to Bonnie Schwartz for innumerable discussions relating to this study. I would also like to thank Ianthi Tsimpli, Kees de Bot, Laurie Stowe, Martha Young-Scholten and Rex Sprouse for discussion of various aspects of this article. For helpful feedback, I thank audiences at the Universities of Durham, Hawai'i, Tübingen and Groningen as well as at EUROSLA 12 in Basel. Finally, I am grateful to two anonymous *Second Language Research* reviewers for instructive criticisms. All remaining errors are my responsibility.

VIII References

Abraham, W. 1986: Word order in the middle field of the German sentence. In Abraham, W. and de Meij, S., editors, *Topic, focus and configurationality*. Amsterdam: John Benjamins, 14–38.

Bley-Vroman, R.W. 1990: What is the logical problem of foreign language learning? In Schachter, J. and Gass, S., editors, *Linguistic perspectives on second language acquisition*. Cambridge: Cambridge University Press, 41–68.

—— 1997: Features and patterns in foreign language learning. Unpublished manuscript, University of Hawaii, Manoa.

Bornkessel, I., Schlesewsky, M. and **Friederici, A.** 2002: Grammar overrides frequency: evidence from the on-line processing of flexible word order. *Cognition* 8, B21–30.

Chomsky, N. 1981: *Lectures on government and binding*. Dordrecht: Foris.

—— 1986: *Knowledge of language: its nature, origin and use*. New York: Praeger.

—— 1995: *The minimalist program*. Cambridge, MA: MIT Press.

Clahsen, H. and **Muysken, P.** 1989: The UG paradox in L2 acquisition. *Second Language Research* 5, 1–29.

Coppieters, R. 1987: Competence differences between native and near-native speakers. *Language* 63, 544–73.

Dekydtspotter, L., Sprouse, R.A. and **Anderson, B.** 1997: The interpretive interface in L2 acquisition: the process-result distinction in English-French interlanguage grammars. *Language Acquisition* 6, 297–332.

den Besten, H. and **Webelhuth, G.** 1990: Stranding. In Grewendorf, G. and Sternefeld, W., editors, *Scrambling and barriers*. Amsterdam: John Benjamins, 77–92.

Ellis, N. 2002: Frequency effects in language processing: a review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition* 24, 143–88.

Epstein, S.D., Flynn, S. and **Martohardjono, G.** 1996: Second language acquisition: theoretical and experimental issues in contemporary research. *Behavioral and Brain Sciences* 19, 677–758.

Grewendorf, G. and **Sabel, J.** 1994: Long scrambling and incorporation. *Linguistic Inquiry* 24, 263–308.

—— 1999: Scrambling in German and Japanese: adjunction versus multiple specifiers. *Natural Language and Linguistic Theory* 17, 1–65.

Hale, K. 1996: Can UG and the L1 be distinguished in L2 acquisition? *Behavioral and Brain Sciences* 19, 728–30.

Hawkins, R. and **Chan, Y.** 1997: The partial availability of universal grammar in second language acquisition: the 'failed functional features hypothesis'. *Second Language Research* 13, 185–211.

Hoberg, U. 1981: *Die Wortstellung in der geschriebenen deutschen Gegenwartssprache [Word order in contemporary written German]*. München: Hueber.

Höhle, T. 1982: Explikation für 'normale Betonung' und 'normale Wortstellung' [Explication of 'normal intonation' and 'normal word order']. In Abraham, W., editor, *Satzglieder im Deutschen [Sentential constituents in German]*. Tübingen: Narr, 75–153.

Hopp, H. 2002: Constraints on word order variation: learnability and UG in advanced English–German and Japanese–German interlanguage. Unpublished MA dissertation, University of Durham.

—— 2004: Syntactic and interface knowledge in advanced and near-native interlanguage grammars. In Foster-Cohen, S., Sharwood-Smith, M.,

Sorace, A. and Ota, M., editors, *EUROSLA yearbook 2004*. Amsterdam: John Benjamins, 67–94.

—— in progress: Grammatical knowledge and its interfaces in near-native interlanguage grammars. Unpublished manuscript, University of Groningen.

Ishihara, S. 2001: Stress, focus and scrambling in Japanese. In Guerzoni, E. and Matushansky, O., editors, *A view from building E39*. MIT Working Papers in Linguistics 39. Cambridge, MA: MIT Press, 151–86.

Koizumi, M. 1995: Phrase structure in minimalist syntax. Unpublished PhD dissertation, MIT, Cambridge, MA.

Lasnik, H. and **Saito, M.** 1992: *Move* α. Cambridge, MA: MIT Press.

Lenerz, J. 1977: *Zur Abfolge nominaler Satzglieder im Deutschen [On the sequence of nominal constituents in German]*. Tübingen: Narr.

Liceras, J. 1993: Los juicios de gramaticalidad en terreno movedizo [Grammaticality judgements in unclear situations]. *Foro Hispànico* 6, 15–29.

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.

—— 2000: Revisiting Universal Grammar. *Arbeiten zur Mehrsprachigkeit [Studies in multilingualism]* SFB 538, B (1), 1–23.

Müller, G. 1996: A constraint on remnant movement. *Natural Language and Linguistic Theory* 14, 355–407.

—— 1998: *Incomplete category fronting*. Dordrecht: Kluwer.

—— 1999: Optimality, markedness and word order in German. *Linguistics* 37, 777–818.

Müller, G. and **Sternefeld, W.** 1993: Improper movement and unambiguous binding. *Linguistic Inquiry* 24, 461–507.

Oka, T. 1996: Scrambling in Japanese and English. *Formal approaches to Japanese linguistics I*. MIT working papers in linguistics 29, 361–88.

Papp, S. 2000: Stable and developmental optionality in native and non-native Hungarian grammars. *Second Language Research* 16, 173–200.

Pinker, S. 1984: *Language learnability and language development*. Cambridge, MA: MIT Press.

Prévost, P. 1999: The second language acquisition of the Split CP structure. In Klein, E.C. and Martohardjono, G., editors, *The development of second language grammars: a generative approach*. Amsterdam: John Benjamins, 81–108.

Reis, M. 1995: Extractions from verb-second clauses in German? In Lutz, U. and Pafel, J., editors, *On extraction and extraposition in German*. Amsterdam: John Benjamins, 45–88.

Rochemont, M. 1989: Topic islands and the subjacency parameter. *Canadian Journal of Linguistics* 34, 145–70.

Saito, M. 1985: Some asymmetries in Japanese and their theoretical implications. Unpublished PhD dissertation, MIT, Cambridge, MA.

—— 1989: Scrambling as semantically vacuous A'-movement. In Baltin, M. and Kroch, A., editors, *Alternative conceptions of phrase structure*. Chicago, IL: University of Chicago Press, 182–200.

Sauerland, U. 1999: Erasability and interpretation. *Syntax* 2, 161–88.

Schachter, J. 1996: Maturation and the issue of universal grammar in L2 acquisition. In Ritchie, W. and Bhatia, T.K., editors, *Handbook of second language acquisition*. New York: Academic Press, 159–93.

Schlesewsky, M., Fanselow, G., Kliegl, R. and **Krems, J.** 2000: The subject preference in the processing of locally ambiguous wh-questions in German. In Hemforth, B. and Konieczny, L., editors, *German sentence processing*. Dordrecht: Kluwer, 65–94.

Schreiber, T. and **Sprouse, R.A.** 1998: Knowledge of topicalization and scrambling in English-German interlanguage. *McGill Working Papers in Linguistics* 13, 162–72.

Schwartz, B.D. and **Sprouse, R.A.** 1996: L2 cognitive states and the full transfer/full access model. *Second Language Research* 12, 40–72.

—— 2000: When syntactic theories evolve: consequences for L2 acquisition research. In Archibald, J., editor, *Second language acquisition and linguistic theory*. Oxford: Blackwell, 156–87.

Takahashi, D. 1993: Movement of wh-phrases in Japanese. *Natural Language and Linguistic Theory* 11, 655–78.

Takano, Y. 2000: Illicit remnant movement: an argument for feature-driven movement. *Linguistic Inquiry* 31, 141–56.

Tsimpli, I-M. and **A. Roussou** 1991: Parameter resetting in the L2. *UCL Working Papers in Linguistics* 3, 149–69.

Tsujioka, T. 2001: Improper remnant movement. *Proceedings of NELS* 31, 483–500.

Uszkoreit, H. 1986: Constraints on order. *Linguistics* 24, 883–906.

Vikner, S. and **Schwartz, B.D.** 1996: The verb always leaves IP in V2 clauses. In Belletti, A. and Rizzi, L., editors, *Parameters and functional heads*. Oxford: Oxford University Press, 11–62

White, L. 2003: *Second language acquisition and universal grammar*. Cambridge: Cambridge University Press.