BA Thesis English Language and Culture, Utrecht University.
Samara van der Burgh
3696677
Dr. W.C.H. Philip and prof. dr. R. Kager
February 2013

Second Language Acquisition in Optimality Theory:
Intralingual Errors

Introduction

This paper discusses current studies related to the second language acquisition of syntax within the framework of Optimality Theory. First there is an explanation of what Optimality Theory is, as originally proposed by Tesar and Smolensky (1998), and how language learning functions in this theory. Since Optimality Theory was originally designed to explain phonetic features of languages, there is a brief discussion on how this theory is applicable to syntactic features of a language. The question under discussion is whether or not an intralingual error can best be attributed to Universal Grammar and, more specifically, whether or not such errors are best discussed within an Optimality Theory framework. An intralingual error is an error that cannot be attributed to either first language influence or to exposure to the target language. The main claim is that in an Optimality Theory framework there is one kind of intralingual error that can be analysed as an emergence of the unmarked. The emergence of the unmarked occurs in an early stage of the interlanguage grammar because markedness constraints are ranked above the constraints of the first language. The reason that second language learners adopt a grammar in which the unmarked structure can emerge is to reduce the burden on the working memory.

Optimality theory: a theoretical framework

Optimality Theory (OT) is based upon the assumption that Universal Grammar (UG) consists of a set of constraints (Con) that are universal and a universal function, known as Gen, which is able to generate all possible candidates[1] of an underlying input (Tesar and Smolensky, 1998). In OT there are sets of multiple constraints which need to be ranked in the correct order for a given language. What distinguishes one language from another are differences in these constraint

---

[1] In syntax Gen is restricted by X-bar theory and as a result Gen cannot create every possible candidate. The candidates that Gen creates have to conform to the rules of X-bar theory.

rankings. In other words, constraints are arranged in a dominance hierarchy. For example in (1) the constraint C_1 dominates C_2 and C_2 dominates C_3, so the dominance hierarchy can be represented as: $C_1 \gg C_2 \gg C_3$, in which C_1 is the most dominant constraint. In a tableau, which shows how a constraint ranking applies to the analysis of a given input during comprehension, an exclamation mark indicates which candidate is considered incorrect.

(1)

Candidates	C_1	C_2	C_3
a) CVC	*!		
☞ b) CV		*	*

In (1) the underlying input consists of phonemes and these consist of consonants (C) and vowels (V). On the basis of the underlying input Gen generates the possible candidates, for example, a sequence of phonemes, such as /bɪt/ (1a) and /bɪ/ (1b). The optimal output is one of the generated candidates. For syntax the input consists of phrase structures, which consist of lexical items and functional categories (Legendre, 2000). The process of producing the correct output is represented in (2).

(2) Input → | GEN | → set of possible candidates → | EVAL | → optimal output
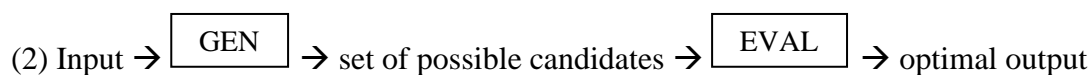
In EVAL (evaluation) the constraints are checked in order to determine the optimal candidate. In the case of (1), for example, the optimal candidate is (1b), which is indicated with ☞, even though it violates more constraints than candidate (1a). This is possible because "the grammatical candidate minimally violates the constraints, relative to the constraint ranking" (Tesar and Smolensky, 1998). For example, the optimal output (1b) does not violate the more highly ranked constraint C_1, though it does violate C_2 and C_3. Characteristic of all constraints is that they are violable; if a constraint is violated then it gets a mark, signified by a *, and the least marked, with

respect to the dominance hierarchy, is the optimal form. Another property of constraints is that they can be in conflict. Constraints can be in conflict because there are two general types of constraints: Markedness constraints, which regulate the well-formedness of the output, and faithfulness constraints that keep the input the same as the output (Prince and Smolensky, 1993). If the optimal candidate differs in some way from the input, a faithfulness constraint is being violated. As we will see below, an example of this is when one constraint forces movement (C_1), while another constraint prohibits movement (C_2). Depending on the constraint ranking in the dominance hierarchy a language can allow either a marked or an unmarked structure. If faithfulness constraints dominate the markedness constraints the optimal output is a marked structure. On the other hand, if the markedness constraints dominate the faithfulness constraints, this is considered to be an unmarked structure.

By a process known as constraint demotion, the highest ranked constraint that is violated by the winner is demoted down in the hierarchy and as a result another candidate becomes the optimal output. Error Driven Constraint Demotion (EDCD) is the process by which the correct ranking is achieved and is therefore a crucial part of the learnability of language in OT. If the target hierarchy is $C_2 \gg C_3 \gg C_1$, and the learner has adopted an incorrect hierarchy as in (1), then the error is detected and the highest ranked constraint that is violated by the loser (1b) is demoted, which leads to a new hierarchy and another optimal candidate, as is shown in (3).

(3)

Candidates	C_2	C_3	C_1
☞ a) CVC			*
b) CV	*!	*	

Since there are always multiple possible candidates, a tableau always contains a winner (1a), which is the optimal form in the target language, and also a loser (1b). However, there is always one optimal output while there are multiple losers, and in order to construct an informative winner/loser pair the correct loser has to be selected by the learner. This results in a process where nothing in the hierarchy changes when the optimal form is the correct form but if the optimal form is incorrect constraint demotion takes place. Eventually this will lead to a total ranking that corresponds with a target language. A distinction is made between a stratified hierarchy, where there are multiple constraints in a stratum, for example, $(C_1, C_2, C_3) >> (C_4, C_5, C_6) \dots >> (C_{10}, C_{11}, C_{12})$, and a total ranking, where each stratum contains only one constraint $(C_1 >> C_2 >> \dots >> C_8)$. The tableau in (4) represents a stratified hierarchy, namely, $(C_1, C_2, C_3) >> (C_4, C_5, C_6)$. The dotted lines indicate that the constraints are in one stratum. The tableau in (4) shows that in a stratified hierarchy there is only one optimal output (4a), which is similar to the totally ranked hierarchy in (3).

(4)

	C_1	C_2	C_3	C_4	C_5	C_6
☞ A						*
B	*			*!	*	
C		*	*!			

Problematic for all learnability theories is the language learning paradox: a grammar is necessary to perceive the underlying structure of an overt form but the grammar can only be acquired if the underlying structure is known. According to Tesar and Smolensky (1998) the solution to this paradox is that the learner starts with a randomly chosen hierarchy and subsequently the learning algorithm EDCD applies. Every time an error occurs, the constraint will be demoted and a new hierarchy will be adopted and this process continues until the correct hierarchy, which is to say

the correct grammar, is attained. According to Tesar and Smolensky (1998) an error can be detected on the basis of informative evidence. This can be either positive evidence or implicit negative evidence. Positive evidence is information from the environment of the learner; on the basis of utterances that are heard in the environment a learner knows which constructions are allowed in that language. Implicit negative evidence is provided by the competing candidates in a hierarchy. If a winner/loser pair is selected the selected loser provides the learner with implicit negative evidence.

Optimality Theory: Syntax

Originally OT was a theory that was mainly used in phonology but research has shown that it can also be applied to other types of linguistic forms, such as semantic and syntactic representations. Some examples below will explain how OT can be applied to syntax and show how syntax is accounted for within an OT framework.

The idea that all languages are similar but differ only in their surface form is an aspect of languages that is difficult to explain. According to Legendre (2000), a theory of UG based upon the idea that there are dichotomous parameter settings and principles that are inviolable, cannot explain the complexity of the surface forms of languages. The problem arises because a principle is inviolable and as a result all languages have to obey to such a principle. However, this does not capture the facts in all cases. In German, for example, this Principle and Parameter (P&P) theory does not apply because for some types of structures overt subjects are obligatory (5a) but for other types null subjects are used (5b). Within a P&P framework there is an Extended Projection Principle (EPP) that states that every clause must have an overt subject (Legendre, 2000). However, in German the EPP is satisfied in (5a) and (5c) but not in (5b) and this differs from which candidates are actually grammatical (5ab).

(5) a. Es wurde   schön   getanzt   'es' is the subject

    It   was   beautifully danced

  b. Schön wurde getanzt.   No subject

  c. * Schön wurde es getanzt.   'es' is the subject

Because in a Principle and Parameters framework constraints are inviolable, Legendre (2000)

argues for adopting OT in order to explain syntactical differences in languages. OT allows

constraints to be violated, and as a result there is constraint interaction. In the German example

sentence (5a) satisfies the EPP but sentence (5b) violates it. This would not be possible if the EPP

has to be satisfied. However, if there is a constraint that has an effect similar to that caused by the

EPP and if this constraint is dominated by another constraint, then this can explain why (5a) and

(5b) are both allowed, even though (5b) violates the EPP and (5a) does not.

Since OT was originally designed to explain phonetic aspects of languages, it is not until

recently that the theory has been applied to syntax. For example, Grimshaw (1997) explains the

subject-verb inversion that occurs due to wh-movement within an OT framework. In a declarative

sentence, such as (7), the subject precedes the auxiliary, while in an interrogative phrase (6) the

auxiliary precedes the subject.

(6) What will he read?

(7) He will read books.

Grimshaw (1997) argues that there are three constraints that account for this inversion. The first

constraint is OP-SPEC, which specifies that "syntactic operators must be in specifier position"

(Grimshaw, 1997), the second constraint is OB-HD, which dictates that a projection always has to

have an audible head, and the third constraint is STAY, which prohibits movement. The tableau in

(8) shows the correct ranking for English interrogatives. Sentence (8b) is ungrammatical[2] because *He will read what* violates OP-SPEC; '*what*' is a syntactic operator and in order to satisfy OP-SPEC it has to be in specifier position, which it is not. The specifier position is unavailable because '*he*' is in specifier position. Another ungrammatical candidate is (8c) because '*what he will read*' violates OB-HD. According to this constraint every projection has a head, which means that every CP and VP has a head. In the sentence *what he will read* (8c) OB-HD is violated because in order for the CP to have a head subject-auxiliary inversion has to occur. The most optimal parse is (8a) even though it violates STAY twice; once because of the movement of '*what*' and once more because of the movement of '*will*' to the head position of CP.

(8)

Candidates	OP-SPEC	OB-HD	STAY
a) ☞ [$_{CP}$ what$_j$ will$_i$ [$_{IP}$DP e$_i$ [$_{VP}$ read t$_j$]]]			**
b) [$_{IP}$ DP will [$_{VP}$ read what]]	*!		
c) [$_{CP}$ what$_j$ e [$_{IP}$ DP will [$_{VP}$ read t$_j$]]]		*!	*

Even more important is that the same constraints in the same ranking can select the correct parse for the declarative sentence (7). The sentence '*will he read books*' violates STAY but the declarative sentence (7) violates none of the constraints, and is for that reason the optimal candidate. According to Grimshaw (1997) candidates that are generated by Gen are restricted by X-bar theory and all candidates must be extended projections. Because of these restrictions, Gen cannot generate an infinite number of candidates.

On the basis of the constraints and their ranking as proposed by Grimshaw (1997), it can be concluded that OT is not only a theory that is useful in the field of phonology but that it is also

---

[2] In a specific context sentence (8b) is grammatical. For example, when someone asks: 'He will read a book' and the response is: 'He will read what?'. In this case the sentence 'he will read what' is an echo-question and is grammatical. However, in the case of (8b) as it is presented in this study the sentence is a matrix interrogative and not an echo-question. As a result sentence (8b) is considered ungrammatical.

capable of explaining the syntactic features of language. Movement, which is traditionally explained within a Principle and Parameters framework, has been accounted for by dominance hierarchies containing constraints such as STAY.

Optimality Theory: Second Language Acquisition

The Emergence of the Unmarked

Related to second language acquisition (L2A) is the emergence of the unmarked[3].This phenomenon is relevant because it can explain errors that are made by second language learners. The main idea behind the emergence of the unmarked is that if a language has a marked structure it also has the unmarked structure (McCarthy and Prince, 1994). The unmarked structure emerges if a marked structure, that is generally allowed, is banned in a particular situation. If in the constraint hierarchy a markedness constraint is dominated by a faithfulness constraint the optimal candidate that is selected is a marked structure. An example of such a hierarchy is: $C_{1 \text{ Faithful}} \gg C_{\text{Marked}} \gg C_{2 \text{ Faithful}}$. The effects of C_{Marked} are invisible because they are blocked by a faithfulness constraint (C_1). However, if in a specific context $C_{1 \text{ Faithful}}$ is not relevant, C_{Marked} becomes active and this situation is called the emergence of the unmarked.

A clear example of the emergence of the unmarked comes from Wakashan (Becker, 2011). In this language the dominance hierarchy for syllables is as follows: IO-MAX (9) >> NOCODA (10) >> BR-MAX (11).

(9) IO-MAX: Every segment of the input corresponds with the output (no deletion)

(10) NOCODA: Syllables do not end with a consonant

(11) BR-MAX: Every segment of the base appears in the reduplicant (no deletion)

---

[3] The emergence of the unmarked is a term first used by McCarthy and Prince (1994).

This ranking of constraints follows the pattern of $C_{1\ Faithful} \gg C_{Marked} \gg C_{2\ Faithful}$. In the tableau (12ab) it can be seen that IO-MAX is the constraint that rules out (12b) because this candidate differs from the input.

(12)

Input: /tʃims-'i:h/	IO-MAX	NOCODA	BR-MAX
☞ a.) tʃim . s'i:h		**	
b.) tʃi . s'i:	**!		
Input: / tʃims-'i:h tʃims-'i:h/	IO-MAX	NOCODA	BR-MAX
☞ c.) tʃi. tʃim . s'i:h		**	****
d.) tʃim. tʃim. s'i:h		***!	***

However, in the case of the reduplicant (12cd) IO-MAX is irrelevant. This is because reduplicants are in correspondence with their base and this is checked by the faithfulness constraint BR-MAX. This differs from the situation in (12ab) where there is a correspondence between the input and output, which is checked by IO-MAX. As a result the markedness constraint NOCODA emerges and rules out candidate (12b) because it has more violations than candidate (12a).

For L2A the emergence of the unmarked is crucial because it can occur in the IL grammar, which is a stage where the learner has not yet fully acquired the TL. Broselow (1998) argues that, in L2A of phonology, it is the unmarked form that emerges in the IL grammar. Because of a new ranking of constraints markedness effects that are not visible in the L1 or TL become visible in the IL grammar. An example of this is the IL grammar of Chinese speakers (L1) of English (L2). In the dominance hierarchies of the L1 and the L2, the constraint NO VOICED OBS CODA (13) is dominated by NO OBS CODA (14), which means that the effects of (13) are invisible.

(13) NO VOICED OBS CODA: Syllable codas may not contain a voiced obstruent

(14) NO OBS CODA: Syllable codas may not contain an obstruent

The ranking in English is: MAX (C), DEP (V), IDENT (VOI) >> NO VOICED OBS CODA, NO OBS CODA. The ranking in Chinese is: NO OBS CODA, NO VOICED OBS CODA >> MAX (C), DEP (V), IDENT (VOI). However, the L2 learners apply devoicing, which can only occur if the constraint (13) is ranked high in the hierarchy. Tableau (15) shows the ranking in the IL grammar.

(15)

/vIg/	NO VOICED OBS CODA	MAX (C), MAX DEP (V)	IDENT (VOI)	NO OBS CODA
☞ a.) vIk			*	*
b.) vIg	*!			
c.) vI		*!		

L2 learners produced /vIk/ and this means that the coda from the input is devoiced. Devoicing occurs because the constraint NO VOICED OBS CODA needs to be satisfied because it is highly ranked. In order for the effects of this constraint (13) to appear, it must be assumed that this constraint is also present in the grammar of both English and Chinese, and that it is due to a new constraint ranking in the IL grammar that the effects of (13) become visible. In this IL grammar (15), the effects of NO VOICED OBS CODA are observed and this is the emergence of the unmarked in the IL grammar. The same effect can be expected to occur in the IL grammar with respect to syntax. A possible reason, which will be discussed in more detail, that explains why a L2 learner adopts a constraint ranking that is neither that of the L1 nor that of the L2 is that this is done to reduce the burden on the working memory.

Second language acquisition: Intralingual Errors

To be able to understand L2A from an OT point of view, it is crucial to discuss IL grammars. An IL grammar is a grammar at a stage between the L1grammar and the target L2 grammar and in this stage the speaker is not yet fully proficient in the TL. Second language users produce errors when they are learning a L2 and the type of error that is produced is important because it can explain how L2A functions. There are two types of errors: interlingual and intralingual errors. An interlingual error is caused by L1 interference and an intralingual error is a type of error that occurs because of exposure to the L2. This means that in an attempt to adopt the correct target grammar the L2 learner makes an error. In an OT framework intralingual errors can be explained as an emergence of the unmarked.

Research by Hyltenstam (1994) on resumptive pronouns in post-nominal relative clauses showed that subjects with different L1's, namely Finnish, Spanish, Greek and Persian, all produced resumptive pronouns, even though the TL was Swedish and this language does not allow resumptive pronouns. Since Greek and Persian both allow resumptive pronouns, the use of them in the IL grammar could be due to L1 transfer. However, Finnish and Spanish both lack the presence of resumptive pronouns in relative clauses and therefore the emergence of this phenomenon in the IL grammar cannot be explained by transfer. According to Eckman (2004) these errors can be explained by the emergence of the unmarked. The constraint hierarchies for the L1s that are proposed by Eckman (2004) are presented in (16). Spanish, Finnish and Swedish have the same ranking (16a), Greek is represented in (16b) and Persian in (16c).

(16)   a. OP SPEC >> STAY >> DEP >> *OBL GAP >> *IO GAP >> DO GAP >> SU GAP

        b. *OBL GAP >> *IO GAP >> OP SPEC >> STAY >> DEP >> *DO GAP >> *SU GAP

        c. *OBL GAP >> *IO GAP >> *DO GAP >> OP SPEC >> STAY >> DEP >> *SU GAP

The GAP-constraints, which are markedness constraints, penalise structures with a gap (17a-d). All the sentences in (17a-d) contain gaps and depending on the gap-position each sentence violates one of the GAP-constraints. For example, (17a) violates OBL GAP because there is a gap in the adposition, (17b) violates IO GAP because there is a gap in the indirect object position, (17c) violates DO GAP because there is a gap in the direct object position and (17d) violates SU GAP because there is a gap in the subject position. In a language, such as Swedish (16a), the GAP-constraints are ranked low and as a result this language allows gaps (17a-d). However, if in a language, such as Greek (16b) or Persian (16c), the GAP-constraints are ranked high, these constraints must be satisfied. In order to satisfy a GAP-constraint the antecedent of the matrix sentence can relativized. A possible relativization strategy is the insertion of a resumptive pronoun in the gap-position. For example, if in a language OBL GAP is the highest ranked constraint and therefore must be satisfied, a relativization strategy that can be applied is to insert a resumptive pronoun in the gap-position (18a-d). In both Greek and Persian this type of relativization strategy is used to satisfy the GAP-constraints.

(17)   a. There is the woman who I read about ___ in an article. (OBL GAP)

        b. There is the woman to who I sent ___ an application. (IO GAP)

        c. There is the woman who I saw ___. (DO GAP)

        d. There is the woman who ___ is my sister. (SU GAP)

(18)   a. There is the woman who I read about *her* in an article. (OBL GAP)

        b. There is the woman to who I sent *her* an application. (IO GAP)

        c. There is the woman who I saw *her*. (DO GAP)

        d. There is the woman who *she* is my sister. (SU GAP)

For example, in Greek, which is a language that does not allow a gap in the indirect object position, the sentence in (19) is grammatical. The resumptive pronoun *'tis'* is inserted to fill the

gap that is in indirect object position and as a result the constraint IO GAP is satisfied. Languages

such as Swedish and Spanish, do not allow this type of structure (19).

(19)   I   ghata pu   **tis** edosa to   gala

      The cat   that her I gave the milk

      'The cat that I gave (her) milk'

The constraint DEP entails that "every element in the output stands in correspondence with the

input" (Eckman, 2004). As discussed before, markedness constraints are generally ranked higher

than faithfulness constraints and Eckman (2004) proposes that in an early stage of IL grammar

markedness constraints are ranked above the L1 constraints. This means that the set of

markedness constraints penalizes constructions that have a gap in the relative cause. The

hierarchy in the IL grammar can be represented as: $C_{Marked} \gg C_{L1}$. Because the L1 constraint

hierarchy is also part of the IL grammar it allows L1 transfer. The motivation for the proposal

that markedness constraints are ranked higher than the L1 constraints is based upon the

assumption that an IL grammar is different from a L1 grammar. This is based on the observation

that errors that are made by children in L1A differ from errors that are made by L2 learners. From

this follows the conclusion that the hierarchy in the L2 cannot be the same as the L1 hierarchy.

The proposal is further motivated by the fact that, on the hand, L2 learners exhibit transfer, and

on the other hand, L2 learners have problems with learning structures in the TL, even though this

structure exists in their L1. If the IL grammar is the same as the L1 grammar, such problems

should not occur. Finally, Eckman (2004) argues that complex structures are difficult to process

for L2 learners and in this case the sentences in (17a-d) are more complex than the sentences in

(18a-d) because a gap is more difficult to process. As a result a learner adopts a hierarchy in

which the optimal output is a sentence that is relatively easy to process.

If Eckman's (2004) proposal about the dominance hierarchy in the IL grammar is correct, the IL ranking, with Swedish as TL, and either Spanish or Finnish as L1, will be as in (20). In (20) the faithfulness constraints of the L1 are intact (OP SPEC >> STAY >> DEP) but they are dominated by the markedness constraints, which are the GAP-constraints. This situation is similar to the situation in which an unmarked structure can emerge; in a particular situation, which in this case is the IL grammar, markedness constraints dominate faithfulness constraints.

(20) *OBL GAP >> *IO GAP >> *DO GAP >> *SU GAP >> OP SPEC >> STAY >> DEP

It is precisely this IL ranking that can explain the errors of Hyltenstam's (1984) subjects that cannot be explained by L1 influence or by exposure to the TL. For example, the IL dominance hierarchy in (21)[4] shows that Spanish speakers allow the use of a resumptive pronoun in a DO relative clause type, even though both the L1 and the TL lack a resumptive pronoun in this type of relative clause.

(21)

Candidates	*OBL GAP	*IO GAP	*DO GAP	OP SPEC	STAY	DEP	*SU GAP
a) Bollen jag gav som till pojken				*!			
b) Bollen som jag gav till pojken			*!		*		
c) ☞ Bollen som jag gav _den_ till pojken					*	*	

In this ranking (21) one of the GAP-constraints is already demoted, which could be because learning already has taken place, however, the L1 faithfulness constraints (16a) are still in the same ranking. It is conspicuous that *SU GAP is demoted first because it is not attested in any L1 and therefore the relevance of this constraint is debatable and Eckman (2004) does not provide

---

[4] This example is from Eckman (2004).

evidence for why the constraint *SU GAP should be part of the universal constraint inventory.

Candidate (21a) fatally violates OP SPEC and candidate (21b) fatally violates DO GAP and because

of that candidate (21c) is the optimal candidate. Because a relative pronoun '*den*' is inserted (22),

the candidate (21c) does not fatally violate the constraint DO GAP.

(22)   Bollen som jag gav **den** till pojken

         'Ball   that I   gave it   to   the boy'

This example (21) is a case of the emergence of the unmarked. The sentence in (21c) is an

unmarked structure because the markedness constraints dominate the faithfulness constraints. In

the L1 (Spanish) the markedness constraints are ranked low and thus they have no visible effect

in the L1. However, in the IL grammar a new hierarchy is adopted and as a result the markedness

constraint DO GAP becomes the decisive constraint in order to evaluate which candidate is the

most optimal.

Eckman's research (2004) can be taken as evidence that in L2A of syntax both

markedness constraints and L1 constraints influence the output. It can be concluded that OT is

capable of explaining one type of intralingual error in the IL grammar, which is not caused by L1

interference or exposure to the TL, as an emergence of the unmarked. Intralingual errors derive

from two factors. The first factor relates to UG indirectly: the constraint ranking from the L1

grammar transfers to the IL grammar. The second factor is the presence of highly ranked

markedness constraints in the IL grammar. Eckman (2004) argues that L2 learners adopt this

particular type of dominance hierarchy, namely $C_{Marked} >> C_{L1}$, because it is more difficult to

process a sentence with a gap. However, there is no experimental evidence provided that shows

that this is the reason that the participants adopted this type of IL grammar. The emergence of the

unmarked in the IL grammar provides evidence for the hypothesis that the set of universal

constraints is part of the grammar of a L2 learner, even if these constraints are not visible in the

L1. This means that L1 grammars and IL grammars are the same in the sense they "reflect the […] regularities that are found in the rest of the world's languages" (Eckman, 2004). However, in the interlanguage grammar the universal constraints are only indirectly accessible, namely via the L1.

Second Language Acquisition: The Interlanguage Grammar

Research by Eckman (2004) showed that intralingual errors can be explained within an OT framework as the emergence of the unmarked. However, Eckman (2004) did not provide experimental evidence to explain why a L2 learner adopts a dominance hierarchy in which the L1 constraints are dominated by a set of markedness constraints. According to Eckman (2004) a possible explanation is that such a dominance hierarchy is adopted because a L2 learner wants to simplify a particular structure and does so by shortening the dependency[5]. This means that in L2A the complexity of processing plays an important role.

Further research (Schulz, 2006, 2011: Slavkov, 2009) on intralingual errors provides evidence that the way an IL grammar is structured is influenced by the working memory capacity of a L2 learner. Both Slavkov's (2009) and Schulz's (2006, 2011) research examined wh-questions that were produced by L2 learners of English. In English wh-movement is cyclic and the wh-word is always the specifier of the main clause (23), however, in other languages wh-copying (24) and wh-scope marking (25) can occur (Slavkov, 2009). Both (24) and (25) are examples of medial wh-. Instead of a silent trace, which is grammatical in English, in (24) the wh-word is copied to medial position and in (25) the wh-word appears in the embedded clause, while the whole sentence is marked by the expletive '*what*'.

(23) Who$_i$ do you think t$_i$ John kissed t$_i$?

---

[5] The dependency is shortened by inserting a resumptive pronoun in the gap position.

(24) *Who$_i$ do you think who$_i$ John kissed t$_i$?   (wh-copying)

(25) *What do you think who$_i$ John kissed t$_i$?   (wh-scope marking)

Schulz (2006) carried out three experiments to see whether or not L2 learners allow medial wh-.

In total there were 135 subjects but some subjects were excluded because it was unclear if they

analysed a sentence, such as (26a), as a case of wh-scope marking or as two separate questions

(26b). As a result the data of 85 subjects is used; 44 German, 28 Japanese and 13 English subjects

and the target language was English.

(26) a. What do you think [who we should invite]?

        b. What do you think? *Who we should invite?

The first experiment was an elicited production task in which subjects were presented with a

picture. Each picture had a title that was a question, for example, "this comic strip tells us which

foreign language Calvin should learn" and the participants had to construct the correct question.

In this example the correct question is: "which foreign language do you suggest Calvin should

learn". The results of the first experiment are shown in Table 1.

	Wh-questions	Medial wh-	% medial wh-
English	93	1	1%
German	232	160	69%
Japanese	235	71	30%

Table 1

The first column shows the number of elicited wh-questions and the second column shows the

number of questions that were produced with medial wh-. One English participant allowed

medial wh- once, which is regarded as noise, and this is in line with what is expected because

medial wh- is ungrammatical in English. The German subjects produced medial wh- frequently in

about half of the complex questions and in the Japanese group medial wh- was relatively less frequent; only 30% of the questions contained medial wh-. Of the Japanese subjects 11 subjects consistently produced sentences with medial wh-, while 11 other subjects did not produce any sentence with medial wh-. This suggests that for some language learners medial wh- is allowed, while for others it is not. It must be noted that in Japanese (L1) medial wh- is not allowed. That some Japanese subjects did allow medial wh- cannot be explained by L1 transfer or by exposure to the TL.

The second experiment was an off-line acceptability judgement task and was carried out because the first experiment was inconclusive about whether or not some subjects allow medial wh-. The test consisted of 100 experimental items, of which 21 tokens contained medial wh- and subjects had to judge whether a sentence was acceptable or unacceptable. In accordance with the first experiment the English participants consistently rejected sentences with medial wh-, while 83% of the tokens with medial wh- were accepted by German participants and 35% of the tokens with medial wh- were accepted by Japanese participants. These results are similar to the results of the first experiment and thus it confirms the results of experiment 1. The last experiment was a self-paced reading task, in which the participants were asked to make a stop-decision if the sentence became ungrammatical. Again the results showed that English participants allowed the occurrence of medial wh- only in 15% of the cases, while German participants allowed this 75% of the time and Japanese participants were in-between with 49%.

It must be noted that only a few participants in Schulz's study consistently allowed medial wh- and that there were also subjects that consistently disallowed it. For example in the online acceptance test the results showed that 36% of the Japanese subjects rejected all sentences with medial wh-, while 18% accepted all sentences (Schulz, 2006). The relatively high acceptance rate of medial wh- by German participants could be due the fact that this construction

is allowed in their native language, which means that this is a case of transfer. What is striking about these results is that in Japanese medial wh- is disallowed, which means that when L2 learners allow wh-scope marking in their IL grammar, it cannot be caused by L1 transfer. Because some Japanese participants allow medial wh-, while others reject it, it is unclear if this is a typological phenomenon or that it is language-specific.

There were multiple learner variables taken into account, such as age of onset, amount of class room instruction, time of immersion, working memory capacity and proficiency. Working memory capacity was measured with a reading span task and the proficiency level was measured with a C-test. Both level of proficiency and the capacity of the working memory were variables that influenced the allowance of medial wh- (Schulz, 2006). However, there is a strong correlation between these two variables and later analysis of the data (Schulz, 2011) showed that it is mainly the capacity of the working memory that influences if subjects allow medial wh- in a sentence. According to Schulz (2011) this allowance of medial wh- can be explained by a simplification strategy that learners apply "in order to reduce processing demands"; it is easier to fill the gap (24, 25) than to have a silent trace (23). It is argued that the IL grammar is restructured, not because there is a parse failure, but because there is a processing overload and in order to reduce the processing overload, the grammar is restructured, even though the IL grammar is already adequate for the L2 (Schulz, 2011). Hereby Schulz (2011) supports the proposal by Eckman (2004) that it is more difficult to process non-adjacent dependencies. However, the proposal by Eckman (2004) also entails that the L1 constraints are ranked below the markedness constraints that penalize structures with a gap and on the basis of Schulz's research (2006, 2011) it can be concluded that the reason that a learner adopts such a hierarchy in the IL grammar is to reduce the processing overload.

Research by Slavkov (2009) with English as L2 and Canadian French as L1 provides further evidence for medial-wh in IL grammars. Two experiments were carried out: a grammaticality judgement task and an elicited production task. In the first experiment the participants had to complete a questionnaire, containing questions, such as (27).

(27) What do you think _____ right now?

      a. he is eating   b. is he eating

      c. what he is eating   c. what is he eating

The second experiment was a version of Thornton's (1990) elicited production task and is similar to Schulz's (2006) method. Canadian French is a mixed language in the sense that wh-movement can occur but it can also be in-situ. It is generally assumed that the standard for Canadian French is with medial wh- marking and subject-verb inversion, which is illustrated in (28).

(28)   Qui penses-tu   qui lit   des histoires?

    Who think-you that reads stories? (Who do you think is reading stories)

The results of these experiments showed that in the grammaticality judgment task medial-wh was chosen 15% of the time and in the elicited production task 6% of the wh-questions that were produced contained medial-wh. These findings are similar to the low rate of Japanese speakers of English that allowed wh-scope marking in their IL grammar (Schulz, 2006). Results also showed that higher-proficiency in English corresponded to a lower rate of medial wh- acceptance, thereby implying that proficiency plays a role in the phenomenon of medial wh- in IL grammars. Since French is a mixed language, Slavkov (2009) argues that the long-distance option is preferred by French speakers because it is present in their L1 and it is possible within UG. However, the use of medial wh- in the IL grammar can occur simply because it is "within the hypothesis space of UG", which seems to be the case in Schulz's research (2011).

Slavkov (2009) also examined wh-movement in L2 English with Bulgarian as L1. If indeed UG plays a role in L2A, rather than L1 transfer, then the results should be similar to that of the French speakers. Bulgarian is a language in which long-distance movement is allowed and where in-situ of wh-words does not occur, which is similar to English. So, if wh-copying or wh-scope marking occurs in the IL grammar, this cannot be explained by either the L1 or the TL. Experiments (Slavkov, 2009) showed similar results with Bulgarian as L1 as for French as L1. The Bulgarian subjects accepted medial-wh at a rate of 18% in the grammaticality judgement task and produced medial-wh in 29% of the sentences in the elicited production task. For both the French and Bulgarian subjects the proficiency level had a significant effect on the allowance of medial-wh. Because both speakers with either Bulgarian or French as L1, which are typologically different languages, produced the same effects in the IL grammar, namely medial-wh, even though all speakers had the same TL (English) and these findings combined with similar effects with Japanese as L1, it can be concluded that medial-wh in the IL grammar is indeed a universal phenomenon, rather than a case of simple L1 transfer. The effects of proficiency level on the occurrence of medial wh- (Schulz, 2006; Slavkov, 2009) show that only in an early stage of L2A a learner uses medial wh- as a simplification strategy.

The effect of the working memory capacity on L2A that was found in Schulz's study (2011) is further supported by other research. For example, recent research (Martin et al, 2012; O'Brien, 2006) showed the effects of the working memory on the L2 proficiency level. In an experiment (Martin et al, 2012) subjects had to learn an artificial language and in the test phase the subjects were tested in two conditions: comprehension and production. In the production task the subjects were presented with an English sentence matched with a picture and had to produce the correct translation. In the comprehension task subjects heard an utterance in the foreign language and had to translate the sentence into English. For both conditions there were sentences

with plural nouns which had not been explicitly taught and which therefore could be used to determine if the subjects had deduced the correct grammar. For all subjects there were two within-subjects variables: working memory capacity and phonological short-term memory (PSTM). The working memory entails the storage and processing of information and was measured with a listening span test. PSTM is only storage and was measured by non-word repetition and recognition. Results showed that PSTM and working memory are both good predictors of learning a L2 vocabulary and grammar. The grammar was mostly influenced by the working memory, which is due to the fact that to learn a grammar information needs to be processed and this is done by the working memory, not by the PSTM.

On the basis of research by Schulz (2006, 2011), Martin et al (2012) and Slavkov (2009) the hypothesis that in L2A the complexity of processing plays a role in L2A is confirmed. Both Schulz (2006) and Slavkov (2009) argue that L2 learners use a simplification strategy. On the basis of Eckman's research (2004) it can be expected that the use of medial-wh by L2 learners can be explained as an emergence of the unmarked. However, in order to know of this is indeed the case, the data of Schulz (2006) and Slavkov (2009) should be reconsidered to determine if it is possible to explain the use of medial wh- in the IL grammars in an OT framework, namely, as an emergence of the unmarked. Moreover, an important aspect of these research papers (Eckman, 2004; Schulz, 2006; Slavkov, 2009) is that they provide evidence for the fact that the set of constraints, which are part of UG, are also part of the IL grammar. The research by Schulz (2006, 2011) provides experimental evidence to support the hypothesis that L2 learners adopt a dominance hierarchy in which markedness constraints dominate L1 constraints and that this is done to reduce processing demands. The influence of the working memory capacity on the allowance of medial wh- can be taken as evidence that this hypothesis is correct: if a L2 learner has a big working memory then there is no need to adopt a simplification strategy and as a result

there will be no intralingual error. However, if the working memory capacity is small, the L2 learner needs to simplify the target structure and this is done by adopting an IL grammar in which there are no complex dependency relations.

Conclusion

The emergence of the unmarked has proven to be essential for L2A research. With regards to phonology (Broselow, 1998) as well as syntax (Eckman, 2004) there is experimental evidence that a universal set of constraints is available for L2 learners. This evidence comes from intralingual errors that are made by L2 learners. For example, the use of the resumptive pronoun in relative clauses by L2 learners, while neither the L1 nor the TL allows resumptive pronouns. In an OT framework these intralingual errors can be explained as the emergence of the unmarked and this type of error can only occur because of the availability of universal constraints. Further research (Schulz, 2006, 2011; Slavkov, 2009) has also provided evidence that learners have access to universal information. Similar to other research (Eckman, 2004; Broselow, 1998) experiments showed that L2 learners make errors which cannot be explained by either the L1 or the L2. However, these experiments (Schulz, 2006, 2011; Slavkov, 2009) are not explained within an OT framework and as a result it cannot be confirmed that these errors are an emergence of the unmarked. It is expected that future research will show that these intralingual errors are an emergence of the unmarked and this would confirm that universal constraints are available to L2 learners. On the basis of Eckman's research (2004) it can be concluded that universal constraints are accessible for L2 learners and that the emergence of the unmarked can explain intralingual errors in an OT framework. However, the constraints are only available via the L1, which means that in L2A UG is not directly accessible.

Other research (Schulz, 2006; Martin et al, 2012) has demonstrated the effects of working memory on L2 proficiency. Experiments (Schulz, 2006; Martin et al, 2012) showed that the working memory capacity can explain the variance of individual proficiency levels. Research (Schulz, 2006; Slavkov, 2009) showed that only in an early stage of L2A, when the proficiency level is relatively low, L2 learners make intralingual errors. These errors are made because the L2 learner simplifies the target structure in order to reduce the burden on the working memory. Because of this simplification strategy, the emergence of the unmarked can occur; in order to simplify the target structure a L2 learner adopts a dominance hierarchy in which markedness constraints dominate the set of L1 constraints. To conclude, in L2A acquisition the complete set of universal constraints is available via the L1. The evidence for this claim is the emergence of the unmarked in the IL grammar. The reason that a L2 learner adopts a dominance hierarchy, such as $C_{Marked} \gg C_{L1}$, that allows the unmarked structure to emerge is to reduce the burden on the working memory capacity.

Works Cited

Becker, Michael and Kathryn Flack Potts. "Chapter 58: The Emergence of the Unmarked." *The Blackwell Companion to Phonology*. Ed. Oostendorp, Marc et al. Blackwell Publishing, 2011. p. 1363 – 1279. Print.

Broselow, Ellen et al. "The Emergence of the Unmarked in Second Language Phonology." *Studies in Second Language Acquisition* 20 (1998): 261 – 280.

Eckman, F. "Optimality Theory, Markedness, and Second Language Syntax: The Case of Resumptive Pronouns in Relative Clauses." *Studies in Phonetics, Phonology, Morphology* 10 (2004): 89 – 110.

Grimshaw, Jane. "Projection, Heads, and Optimality." *Linguistic Inquiry* 28.3(1997): 373 – 422.

Hyltenstam, K. "The Use of Typological Markedness Conditions as Predictors in Second Language Acquisition: The case of Pronominal Copies in Relative Clauses." Ed. R. Andersen. *Second Languages: A cross-linguistic perspective*. Rowley: Newbury House Publishers, 1984. p. 39 – 58. Print..

Legendre, Geraldine. "An Introduction to Optimality Theory in Syntax." *Optimality-theoretic syntax*. Ed. Legendre, Geraldine et al. Cambridge: MASS MIT Press, 2000.

Martin, Katherine I. and Nick C. Ellis. "The Roles of Phonological Short-Term Memory and Working Memory in L2 Grammar and Vocabulary Learning." *Studies in Second Language Acquisition* 34 (2012): 379 – 413.

McCarthy, John J. and Alan S. Prince. "The Emergence of the Unmarked. Optimality in Prosodic Morphology." *Papers from the Annual Meeting of the North East Linguistic Society* 24 (1994): 333–379.

O'Brien, Irina et al. "Phonological Memory and Lexical, Narrative, and Grammatical Skills in Second Language Oral Production by Adult Learners." *Applied Psycholinguistics* 27 (2006): 377 – 402.

Prince, Alan and Paul Smolensky. "Optimality Theory: Constraint Interaction in Generative Grammar". Technical Report, Rutgers University Center for Cognitive Science and Computer Science Department, University of Colorado at Boulder (1993).

Schulz, Barbara. "Wh-scope Marking in English Interlanguage Grammars: Transfer and Processing Effects on the Second Language Acquisition of Complex Wh-questions." Diss. University of Hawaii, 2006.

Schulz, Barbara. "Syntactic Creativity in Second Language English: Wh-scope Marking in Japanese-English Interlanguage." *Second language research* 27.3 (2011): 313 – 341.

Slavkov, Nikolay. "The Acquisition of Complex Wh-questions in the L2 English of Canadian French and Bulgarian Speakers: Medial Wh- constructions, Inversion Phenomena, and Avoidance Strategies." Diss. University of Ottowa, 2009.

Tesar, Bruce and Paul Smolensky. "Learnability in Optimality Theory." *Linguistic Inquiry* 29 (1998): 229–268.

Thornton, R. "Adventures in long-distance moving: The acquisition of complex wh-questions." Diss. University of Connecticut, Storrs, 1990.