Foot recursion in interlanguage grammars: A study on expletive infixing

Emma M.P. Haggins and John Archibald
University of Victoria
emma.haggins@gmail.com

Recursion is a fundamental property of generative grammars (Watamull, Hauser, Roberts, & Hornstein, 2014). While there is considerable research on interlanguage grammars (White, 2003; Schwartz & Sprouse, 1996), there has been little research into recursion in interlanguage grammars. To further our knowledge in this area, we aim to answer the question of whether second language learners of English have phonological recursion in their interlanguage grammars. Our focus will be on recursion at the metrical foot level.Using a lexical decision task we conducted a forced-choice selection task. The participants judged the grammaticality of swear words that had been infixed into two different spots in the same word (e.g., fan-fucking-tastic vs. *fantas-fucking-tic), where the first form is well-formed according to the infixing principles of McCarthy (1982) but the second item is ill-formed.We analyzed accuracy data to see (a) whether learners can distinguish well-formed from ill-formed strings, and (b) whether the L1 makes a difference. If participants discriminate between well- and ill-formed strings, it will show that they have recursion in their interlanguage grammar.
Keywords: recursion; interlanguage grammar; foot; phonology; lexical decision task

1 Introduction

Recursion is a fundamental property of generative grammars (Watamull, Hauser, Roberts, & Hornstein, 2014). To further our knowledge in this area, we aim to answer the question of whether second language learners of English have phonological recursion in their interlanguage grammars. Our focus will be on recursion at the metrical foot level.

We conducted a lexical decision task with four L2 speakers of English who all had varying L1s. The L1s studied either had metrical feet, or they did not. The participants then chose between two versions of a word with the only word used as an infix in English, *fucking*, inserted at different locations. For example: fan-fucking-tastic vs. *fantas-fucking-tic. The accuracy of the answers were then analyzed.

Clahsen and Felser (2017) claim that interlanguage grammars have shallow (i.e., limited hierarchical) structure only, this suggets that interlanguage grammars are incapable of representing recursive structures. In this paper, we will be disputing these claims.

2 Literature Review

In this section we will explain the key background information necessary to understand the experiment, including a discussion of the evolution of recursion in human language, the operation Merge, as well as the main properties of recursion. Finally, this section will conclude with an outline of the rest of the paper.

2.1 Evolution

There is a rich literature, which looks at the evolution of language in *homo sapiens* (Berwick, Friederici, Chomsky and Bolhuis, 2013; Berwick and Chomsky, 2016). It is clear that animals have the ability to communicate. Non-human primates have call systems, which are somewhat analogous to words in that different calls have different meanings (e.g. *eagle* versus *leopard*). Species such as whales, or birds have songs, which can have internal structure (e.g., sequences of notes). However, the component parts of these songs (i.e., the notes) do not have individual meanings, so the songs are unlike human sentences. What none of human's common ancestors seem to possess is the generative capacity to produce novel utterances in which constituents are structured recursively from simple representational elements. It is truly a central property of human I-language, and thus, interesting to look at whether we find it in interlanguage grammars.

2.2 Merge

Yang, Crain, Berwick, Chomsky and Bolhuis (2017) discuss the recursive process that is responsible for the formation of linguistic structures. This recursive operation (known as *Merge*) combines two linguistic terms to produce a new, composite term. The new term can then also be merged with another linguistic term, and so on until the phrase is fully derived.

Merge is currently thought to be "the fundamental operation of structure building in human language" (Yang et al., 2017, p. 3). We could then extend this theory and claim that the reason humans developed a language, and other animals didn't, is because we are the only ones that developed Merge (Yang, Crain, Berwick, Chomsky & Bolhuis, 2017).

2.3 Recursion

According to Watumull, Hauser, Roberts and Hornstein (2014) there are three main properties of recursion: computability, induction and boundedness. In this section we will briefly discuss these properties. These are properties which underlie our ability to generate sentences such as *The man in the shirt with stripes wondered whether his outfit was appropriate for the wedding in the cathedral.*

2.3.1 Computability

A finite program of rules, states and symbols, and a mechanism for decoding, encoding and manipulating symbols are all necessary for computation. Using the Turing machine as a good example of computation, Watumull et al. (2014, pg. 1) noted that "the machine generates theorems given inputs by returning intermediate results according to its programmed rules." Thus, the grammar is generative; an infinite set of grammatical sentences can be built via such machinery.

This grammar (what Chomsky calls I-language) is, thus, internalized in the human mind/brain, while E-language (or *external* language – the sentences people actually produce) is generated and constrained by I-language.

They conclude by stating that- "computable functions are therefore those calculable by finite means." The infinite capacity of human language can be modeled computationally.

2.3.2 Induction

Watumull et al. also discuss a key property (first noted by Gödel) of recursive functions: induction.
This property of a generative grammar was originally shown in Post's rewrite rules (Watumull et al., 2014). These rules were in the following form:

φ → ψ ("rewrite φ as ψ")

This type of rule derives hierarchical syntactic structure. So, a verb phrase could be re-written as V + NP. Watumull et al. wrap this up by stating, "a grammar strongly generates hierarchically structured expressions [the I-language] and weakly generates the corresponding strings [the E-language]" (Watumull et al., 2014, p.3). The structure is what conveys grammatical information. The information can then be mapped, via linguistic processing, to the conceptual-intentional (LF) and the sensory-motor (PF) systems.

2.3.3 (Un)boundedness

The final property to discuss is (un)boundedness, which has two important aspects. First, a recursive function may generate an infinite set of possibilities, but only produce a finite output, because of mechanistic constraints. For example, I-language produces E-language. I-language is internalized and it would therefore license every grammatical sentence a human could produce; however, E-language is external, and only illustrates what each human actually says. Not every grammatical sentence gets uttered. No one actually has the time to produce a sentence with 1,000,000,000 words in it.

Secondly, any arbitrarily limited output can be expanded, because recursive functions have no limit (e.g., *The team won the trophy.*; *The coach said that the team won the trophy,*etc.). There is no longest grammatical sentence. Let us turn now to other aspects of this property, such as phonological structure.

2.4 Feet

In most languages, syllables get parsed into metrical feet, which then get parsed into the prosodic word node. Trochaic feet are strong (or prominent) on the left and iambic feet are strong on the right. Three out of the four languages included in this study have foot structure. English, German and Mandarin all have trochaic feet (Weber, 2013; Qu, 2013), whereas French has no foot structure (Özçelik, 2016). In the sections below, we will discuss the structure of German, Mandarin and French with relation to their respective type of foot structure, or the lack thereof.

2.4.1 German foot structure

Figure 1. below shows the foot structure present in German (Weber, 2013). As shown below, the strong syllable (i.e., the more prominent) is on the left of the foot, and the strong foot is on the left of the Word, making the German prosodic structure a trochaic system. In *Figure 1* the symbol ω indicates the prosodic word level, the symbol Σ represents the foot level and the σ indicates the syllable level. The subscript s and w represent strong and weak feet or syllables.

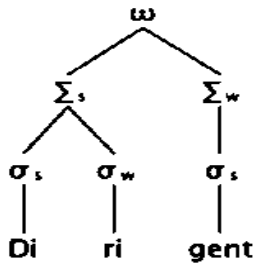

Figure 1. German example of violations of metrical foot structure for the noun *Dirigent* "conductor"

2.4.2 Mandarin foot structure

Below, in Figure 2., is a diagram showing Mandarin foot structure (Qu, 2013). Mandarin feet are also strong on the left (marked with an s on the stronger syllable), making them trochaic as well. In this example PWd is used as a short form for the prosodic word node, the subscript numbers in the Mandarin sentence represent the tones used on each word.

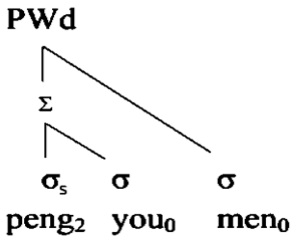

Figure 2. Mandarin foot structure of the phrase *peng$_2$ you$_0$ men$_0$* "friends"

2.4.3 French foot structure

As can be seen in the diagram below, French does not have any foot structure (the syllables are grouped directly into the prosodic word) at all, making it a footless language (Özçelik, 2016). In Figure 3. PPh is used to mark the prosodic phrase node.

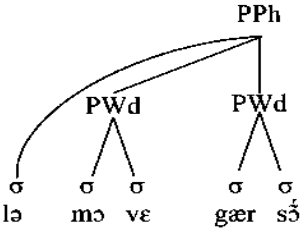

Figure 3. No foot structure is necessary in French for the phrase *le mauvais garçon* "the bad boy"

2.4.4 Foot Structure and Recursion

The tree structures below in Figure 4. show why foot recursion is necessary in English when 'fucking' has been infixed into a word. As can be seen in the diagram below, c. is the only well-formed tree structure in that (a) and (b) show the (illicit) crossing of association lines. This is evidence that foot recursion is necessary (McCarthy, 1982) in order to generate these infixed forms. Note that there is a Foot within a Foot in (c) which is clearly a recursive structure.

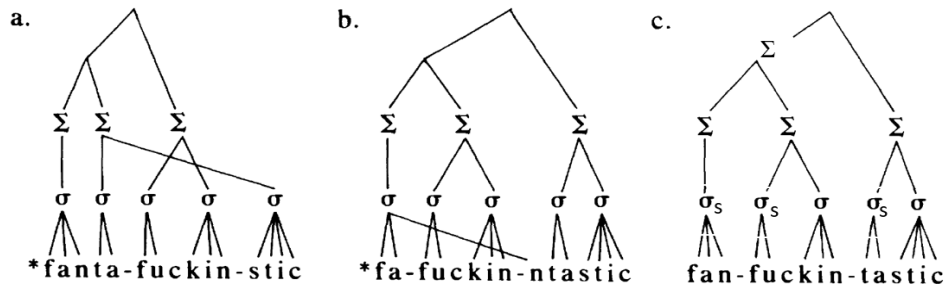

Figure 4. Foot recursion is necessary for infixed forms

2.5 Research Question

This paper aims to answer the question: Do L2 speakers of English have phonological recursion in their interlanguage grammars? Or, as Clahsen & Felser (2017) suggest, are their interlanguage grammars *shallow (*i.e., lacking hierarchical structure)? Using a lexical decision task, L2 learners of English will judge the grammaticality of swear words with infixes in various positions in English words. If participants are accurately able to judge the infixation, it will show that they have recursion (i.e., feet within feet) in their interlanguage grammar. However, if the participants are unable to accurately judge the grammaticality of the infixation, it will show that they do not have recursion in their interlanguage grammar.

3 Methodology

3.1 Participants

There were 5 participants in this study, one 22-year-old native English speaker (as a control), and four-second language speakers. Within the second language speakers there was one female French speaker, one female Mandarin speaker and two German speakers (one male and one female). The second language speakers had been speaking English for a range of 10 to 42 years. They were all of advanced proficiency (all graduate students or professors).

3.2 Materials

The materials used in this experiment were: a background information questionnaire, and a grammaticality judgment task, which was created and run through PsychoPy. The stimuli consisted of 22 pairs of three or four syllable words with the f-word infixed into them, this word list is seen in Appendix A. The pairs consisted of the same word where ‘fucking’ had been infixed into the correct location for one, and an incorrect location for the other. During the task, the incorrect option was listed first 11 times, and the correct option was listed first the other 11 times. It was decided randomly, which one would come first. A

computer and a pair of headphones were also necessary materials for this experiment.

3.3 Procedure

This experiment consisted of two main steps. After signing the HREB approved consent form, participants completed a questionnaire of background information. They then completed the grammaticality judgment task which consisted of the participants listening to and reading the pairs of words, where 'fucking' was either infixed in the grammatical location or the ungrammatical location, and pressing a button on a keyboard that corresponded to their choice.

3.4 Data analysis

Psychopy recorded accuracy and reaction time in an excel spreadsheet. The accuracy of their responses was then analyzed. Reaction time was not analyzed as there were too few participants to allow for statistical analyses.

4 Results

Below are the graphs of the results obtained. Each of the graphs reports on data from a different L1.

The first graph, in Figure 5., is that of the native English speaker control, who had been living in Canada for her entire life (22 years). This participant judged 21 of the 22 pairs of words correctly. The chi squared results for the English speaker were as follows: 1 wrong, 21 correct, p 0.01* $\chi^2(21)$. This validates the experimental task and provides the baseline for our non-native speaker comparison.

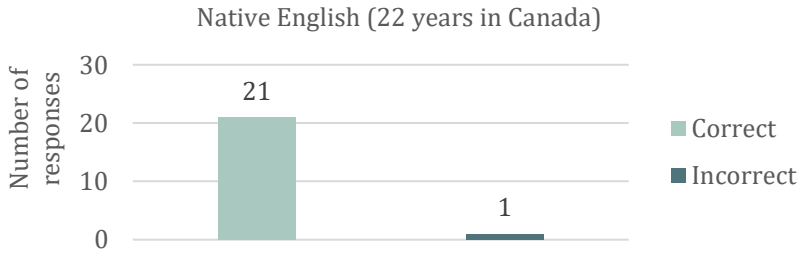

Figure 5. Number of correct and incorrect responses of the L1 English participant

Figure 6. shows the graph representing the responses of the two L1 German speakers. The German L1 participants had been living in Canada for 8 months and 34 years respectively. They both scored correctly on 20 of the 22 pairs of words. The chi squared results for the two German speakers were: 2 wrong, 20 correct, p 0.01* $\chi^2(21)$.

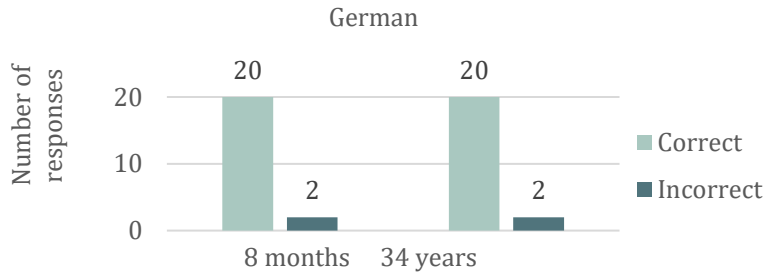

Figure 6. Number of correct and incorrect responses for both L1 German participants

Figure 7. shows the responses of the L1 Mandarin speaker, who had been living in Canada for 3 years. The Mandarin L1 participant got 19 of the pairs of words correct, and 3 incorrect. The chi squared results for the Mandarin speaker was: 3 wrong, 19 correct, p 0.01* $\chi^2(21)$.

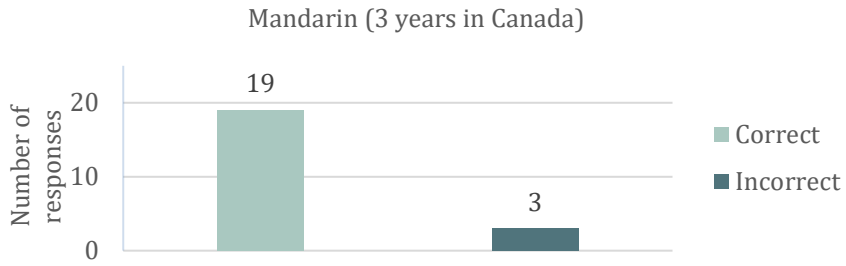

Figure 7. Number of correct and incorrect responses for the L1 Mandarin Speaker

Figure 8. represents the number of correct and incorrect responses for the L1 French speaker who had been living in Canada for 1.5 years. The L1 French speaker scored 18 pairs correctly and 4 pairs incorrectly. The chi squared results for the French speaker was: 4 wrong, 18 correct, p 0.01* $\chi^2(21)$.

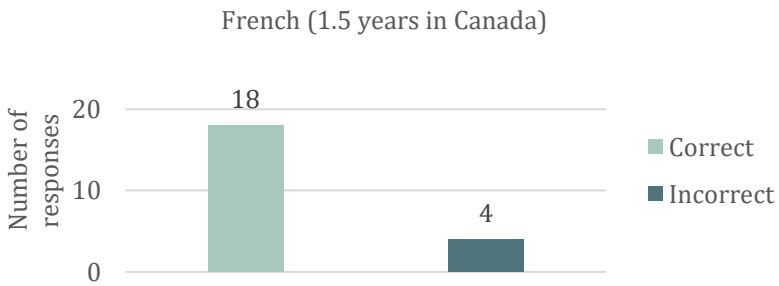

Figure 8. Number of correct and incorrect responses for the L1 French speaker

5 Discussion

The chi-squared test results show that the participants were not guessing when they completed the well-formedness judgment task; the accuracy was well above 50%. We can conclude, therefore, that they are able to accurately judge the grammaticality of expletive infixing, and that this accuracy results from a representation of recursion.

While acknowledging that we do not have enough subjects to make robust cross-linguistic comparisons, the most errors were made by the French speaker, which is consistent with what was expected because French was the only language tested that does not have foot structure. They are having to acquire a new structure in their L2 English.

The female German speaker had been in Canada for the shortest period of time and still performed as well as the male German speaker who had been in Canada for the longest period. This is consistent with hypothesis 2 as well since German foot and stress structure is the most like English. The equal performance of these two subjects shows that this ability is not something found only after lengthy exposure to L2 English.

We should also note that since none of French, Mandarin or German allow any infixing, it is not the case that these speakers are transferring L1 infixing knowledge to make L2 grammaticality judgments; this is something they have acquired in their second language.

6 Conclusion

Given that recursion is a central property of grammar (Watamull, Hauser, Roberts, & Hornstein, 2014), only humans have merge (Yang, Crain, Berwick, Chomsky & Bolhuis, 2017) and humans are usually bilingual, the default assumption should be that interlanguages have recursion, and our data confirm this hypothesis.

The L2 speakers were able to correctly judge the grammaticality of expletive infixing, and whether their L1 had feet seemed to play a part in how accurate they were. This leads to the conclusion that interlanguage grammars have a recursive structure contra the claims of Clahsen and Felser (2017).

Acknowledgements (from the 1st author)

I'd like to thank Dr. John Archibald for his guidance, suggestions and support throughout the last 8 months. I'd also like to thank all of the participants for completing this experiment, and my family and friends for their extra support. I would like to acknowledge that this study was conducted on the traditional territory of the WSÁNEĆ (Saanich), Lkwungen (Songhees), Wyomilth (Esquimalt) peoples of the Coast Salish Nation.

References

Berwick, R. & N. Chomsky (2016). *Why Only Us?* MIT Press.

Berwick, R. C., Friederici, A. D., Chomsky, N., & Bolhuis, J. J. (2013). Evolution, brain, and the nature of language. *Trends in Cognitive Sciences*, *17*(2), 98. https://doi.org/10.1016/j.tics.2012.12.002

Clahsen, H., Felser, C. (2017). Critical commentary: some notes on the shallow structure hypothesis. *Studies in Second Language Acquisition*, 1-14.

Lee, Y., Jang, E., Choi, W. (2018). L2-L1 translation priming effects in a lexical decision task: evidence from low proficient Korean-English bilinguals. *Frontiers in Psychology*, 9(267), 1-10. Doi:10.3389/fpsyg.2018.00267

McCarthy, J. J. (1982). Prosodic Structure and Expletive Infixation. *Language*, *58*(3), 574–590.

Özçelik, Ö (2016). The foot is not an obligatory constituent of the prosodic hierarchy: "stress" in Turkish, French and child English. *The Linguistic Review*, 1-57

Peirce, J., Gray, J.R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., Lindeløv, J.K, (2019) PsychoPy2: experiments in behavior made easy, *Behaviour Research Methods*, 51(1), 195-203

Qu, C. (2013). *Representation and acquisition of the tonal system of Mandarin Chinese*. (PhD thesis). McGill University. Montreal, Canada.

Schreuder, M., Gilbers, D., & Quené, H. (2009). Recursion in phonology. *Lingua*, *119*(9), 1243–1252. https://doi.org/10.1016/j.lingua.2009.02.007

Schürmann, C., Despeyroux, J., & Pfenning, F. (2000). Primitive recursion for higher-order abstract syntax with dependant types. *Theoretical Computer Science*, *266*, 1–57.

Schwartz, B. D., & Sprouse, R. A. (1996). L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research*, *12*(1), 40–72. https://doi.org/10.1177/026765839601200103

Verhagen, A. (2008). Syntax, Recursion, Productivity — A usage-based perspective on the evolution of grammar. *Studies in Slavic and General Linguistics*, *33*, 399–414.

Watumull, J., Hauser, M. D., Roberts, I. G., & Hornstein, N. (2014). On recursion. *Frontiers in Psychology*, *4*, 1–7. https://doi.org/10.3389/fpsyg.2013.01017

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

Weber, S. (2013). *Effects of Lexical Stress Errors on the Intelligibility of German and English*. (PhD) University of Calgary, Canada.

Yang C., Crain S., Berwick R.C., Chomsky N., Bolhuis J.J., The growth of language: Universal Grammar, experience and principles of computation. *Neuroscience and Biobehavioral reviews* (2017)

Appendix A

Stimuli list:

(1)		**Fantastic**	(2)		**Kindergarten**
	a.	Fantas-fucking-tic		a.	Kindergar-fucking-ten
	b.	Fan-fucking-tastic		b.	Kinder-fucking-garten
(3)		**Everybody**	(4)		**Scarborough**
	a.	Every-fucking-body		a.	Scar-fucking-borow
	b.	Everybo-fucking-dy		b.	Scarbo-fucking-row
(5)		**Unbelievable**	(6)		**Irresponsible**
	a.	Unbelieve-fucking-able		a.	Irrespons-fucking-ible
	b.	Unbe-fucking-lievable		b.	Irre-fucking-sponsible
(7)		**Vancouver**	(8)		**Garibaldi**
	a.	Vancou-fucking-ver		a.	Gari-fucking-baldi
	b.	Van-fucking-couver		b.	Garibal-fucking-di
(9)		**Nanaimo**	(10)		**Adventure**
	a.	Na-fucking-naimo		a.	Adven-fucking-ture
	b.	Nanai-fucking-mo		b.	Ad-fucking-venture
(11)		**Winnipeg**	(12)		**Saskatoon**
	a.	Winni-fucking-peg		a.	Saska-fucking-toon
	b.	Wi-fucking-nnipeg		b.	Sa-fucking-skatoon
(13)		**Mississauga**	(14)		**Celebrate**
	a.	Missi-fucking-sauga		a.	Celebra-fucking-te
	b.	Missisau-fucking-ga		b.	Cele-fucking-brate
(15)		**Pollution**	(16)		**Basketball**
	a.	Po-fucking-llution		a.	Basket-fucking-ball
	b.	Pollu-fucking-tion		b.	Bas-fucking-ketball
(17)		**Identical**	(18)		**Information**
	a.	Iden-fucking-tical		a.	Informa-fucking-tion
	b.	I-fucking-dentical		b.	Infor-fucking-mation
(19)		**Watermelon**	(20)		**Burnaby**
	a.	Waterme-fucking-lon		a.	Burna-fucking-by
	b.	Water-fucking-melon		b.	Bur-fucking-naby
(21)		**Abbotsford**	(22)		**Coquitlam**
	a.	Abbots-fucking-ford		a.	Coquit-fucking-lam
	b.	Ab-fucking-botsford		b.	Co-fucking-quitlam