The text and figures in this PDF are approved by the author(s) for publication. Any mistakes in this PDF will not be corrected by the publisher. This PDF was created on 15/06/21.

**Input beyond the Threshold: Explaining Auxiliary Initial Assertions in a British English Early Talker**

**Rebecca Woods, Johannes Heim, and Joel Wallenberg[[1]](#footnote-2)**

**1. Introduction: early word order mappings in child English**

Multiword utterances in infants usually reflect their target variety’s canonical word order for different speech acts astonishingly early: in English-acquiring children, early assertions start with a subject and early polar questions start with an auxiliary (Brown & Bellugi 1964; Gleitman *et al*. 1984). In this paper, we present diary and recorded data from one early talking child acquiring British English, Paddy, who defies this pattern by almost exclusively producing auxiliaries in sentence-initial position in his early production for a period of around 5 months. We argue here that Paddy’s unusual word order results from the interaction between (1) a general, stochastic rule-formation process, (2) a notable and meaningfully large proportion of auxiliary-initial (AuxS) constructions in his input compared with the proportion of auxiliary-medial (SAux) constructions, and (3) precociously early production of multiword utterances. Based on these three factors, Paddy postulates that auxiliaries occur in initial position regardless of the purpose of the utterance and demonstrates this through his production. We claim that his case shows that English word order is not mystically accurately acquired and that it can also serve more specifically as a window into how children map speech acts onto specific word orders.

Paddy is a cognitively typical child growing up in a monolingual English-speaking household in northern England. He is the first of two children, and his two caregivers are both educated to postgraduate level. He began using multiword utterances around 1 year 10 months and has a mean length of utterance by word (MLUw) of more than 3 by age 2;7, indicating that he is linguistically precocious (cf. Brown 1973, Evans and Demuth 2012). His linguistic, specifically syntactic, development is typical in most respects but unusually, he uses auxiliary-initial word order not only for questions, but also for assertions, as exemplified in (1-3). An auxiliary precedes the subject in each example, yet the context reveals that Paddy is not asking questions in these sentences.

(1)[Context: Paddy throws two books straight up into the air and they fall to the floor*.*]
Paddy: Did you [=Paddy][[2]](#footnote-3) throw them – did you catch them in the air! (2;4, diary data)

(2) Friend: Has the dinosaur got little teeth?
Paddy: Has he not. (2;3, diary data)

(3) [Context: Paddy wears a toy as a hat. No-one else has anything on their heads.]
Paddy: Is it on your [=Paddy’s] head! (2;4, diary data)

Using diary data from Paddy aged 2;3-2;5 and audio recordings from 2;3-2;7 as evidence, we propose that Paddy treats the mapping of speech acts onto word order configurations (i.e., clause types) like any other regularization problem: variation in the input—here the distribution of auxiliaries—requires a certain amount of data for patterns to emerge and to be translated into conventionalized mappings. Every infant is confronted with variation in word order in their input, such that if auxiliaries most frequently occur after the subject, the child may treat this as a default word order. Conversely, if auxiliaries most frequently occur in sentence-initial position, the child may treat such a word order as default. In a third case, where auxiliaries are equally frequently in sentence-initial or -medial position, the child may not assume a default position for these elements and will have to seek further evidence for auxiliary placement and its ramifications. If a child assumes a default position, we would expect this mapping to be temporary and hold only until the child encounters sufficient data to map different syntactic configurations onto different illocutionary functions. Before this point, individual deviations from the hypothesized default are treated as lexicalized exceptions.

This proposal is an extension of the classic learning problem in the morphosyntactic domain: form-function mappings have exceptions that need to be recognized as such and the elsewhere case will only materialize once the context-specific rules are identified. We will use Yang’s (2012, 2016 *et* *seq.*) Tolerance Principle model in combination with the Variational Learning model proposed for the clausal domain (Yang 2000, 2002, 2004) to capture a process that is, we claim, much more frequent than is currently reported in the literature. We believe that this process is typically obscured because it occurs in the mind before many children are producing multiword utterances (certainly ones containing auxiliaries), and because auxiliaries generally are rare and often omitted in early child production. The fact that Paddy is an early talker and that his caregiver input exhibits an exceptional asymmetry in the proportion of auxiliary-initial and auxiliary-medial constructions means that we can observe in his production a point where the child’s assumptions about word order do not yet match the adult canonical conventions.

The Tolerance Principle accounts for the regularization problem by positing a proportional frequency of exceptions that are allowed for each rule that the child postulates – that is, how many exceptions will be ‘tolerated’. This proportion constitutes a threshold that, if surpassed, prompts the child to change their postulated rule. Applying this model, we claim that the data in (1-3) result because Paddy first postulates a general AuxS rule and does not initially hear a sufficient number of exceptions to this to prompt a rule change. We compare Paddy’s data with two other early talkers (determined by MLUw before age 3) from the CHILDES database (MacWhinney 2000), Eve (Brown corpus, Brown 1973) and Naima (Providence corpus, Evans and Demuth 2012). This creates, on the face of it, quite a small data set, but none of the other (neurotypical) children in the Brown or Providence corpora have a sufficiently high MLUw before age 3.[[3]](#footnote-4) This suggests that they are not comparable to Paddy in terms of linguistic production at the relevant stage of input saturation and cognitive development. We show that the relative frequency of sentence-initial auxiliaries in Paddy’s caregiver input creates a different context for acquisition than does Eve or Naima’s input. In the remainder of this paper, we compare the input and output data of these children. We then explain how combining two learning mechanisms, Variational Learning and the Tolerance Principle, accounts for Paddy’s exceptional mapping, where Eve and Naima’s input leads them to use only a Variational Learning mechanism. We provide a detailed analysis of the different auxiliaries that Paddy hears and produces. Before we conclude, we spell out implications of our proposal for other clausal instances of Variational Learning and point to diachronic changes underlying the same mechanisms.

**2. Data: Paddy’s exceptional auxiliary-initial assertions**

To further appreciate the exceptional character of Paddy’s word order mappings, consider the questions (a) and assertions (b) in examples (4-6). Eve and Naima map questions and assertions onto AuxS and SAux structures respectively, whereas Paddy uses AuxS for both speech acts.

* 1. a. Eve: Can I have apples, too? (2;1)
	b. Eve: You can have lobster salad (2;2)

(5) a. Naima: Can I have my wallet? (2:9)
 b. Naima: I can have lunch before your ice cream. (2;4)

1. a. Paddy: Can I read that? (2;6, written diary data)
b. Mother: Where is your pen? Oh, there it is!

 Paddy: Can Paddy get it. [Reaches for the pen] (2;3, audio recording)

Paddy’s production data stands out due to the presence of auxiliary-initial assertions. This uncommon speech act mapping goes hand in hand with the fact that these fronted auxiliaries occur before sentence-medial auxiliaries arise, independent of the type of auxiliary. This is unlike what is reported for other English-acquiring children, who typically utter sentence-medial auxiliaries before they are fronted (Stromswold 1990), and for whom inversion of copula BE tends to precede the inversion of all other auxiliaries (Cazden 1972). In other related aspects of his production, however, Paddy does not appear to diverge from the expected path of acquisition. He does not show any irregularities for head directionality (Hirsh-Pasek & Golinkoff 1996). He also clearly distinguishes between auxiliaries and verbs in that he exhibits do-insertion and does not invert lexical verbs (Stromswold 1990), as in (7).

1. [Context: Paddy pushes a toy car over to the other side of the room.]
Paddy: Did you push down. (2;3, written diary data)

He often omits auxiliaries and copula BE altogether (Brown 1973), as in (8).

1. I jumping! I spinning around! I spinning the other way! (2;4, written diary data)

When producing an inflected auxiliary, he also produces an overt nominal subject (Sano and Hyams 1994, Schütze & Wexler 1996), so we do not see utterances such as “Is jumping!”.We should also note here that Paddy ultimately stops using AuxS word order for utterances around 2;8, just after recordings cease. Given Paddy’s generally typical acquisition of English, the question arises as to what motivates Paddy to treat AuxS as the default word order in his early production of auxiliaries.

Table 1 provides an overview of the proportions of structures containing auxiliaries in Paddy’s caregiver input as recorded from age 2;3-2;7, with Eve and Naima’s caregiver input for comparison.[[4]](#footnote-5) Although there are significantly fewer datapoints in recordings from Paddy than in the corpus data from Eve and Naima, the overall trend is clear: the AuxS and SAux constructions in Eve’s and Naima’s input make up similar proportions of the overall input data, whereas there is clear asymmetry in Paddy’s data as SAux constructions are much less frequent.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Paddy** | **Eve** | **Naima** |
| ***Average MLUw (Age)*** | *2.9 (**2;3-2;7)* | *3.23 (1;11-2;3)* | *2.87 (1;3-2;7)* |
| **AuxS** | 116 (12.58%) | 1164 (8.0 %) | 5416 (8.8%) |
| **SAux** | 31 (3.36%) | 953 (6.7%) | 6330 (10.7%) |
| **Overall** | 922 | 14509 | 61695 |

*Table 1: Proportions of AuxS and SAux constructions in caregiver input*

We now take a closer look at Paddy’s production. Here, AuxS clearly outnumber SAux constructions: of the 38 utterances in Paddy’s audio-recorded[[5]](#footnote-6) production which include an auxiliary, 37 contained a sentence-initial auxiliary, and only 1 contained a sentence-medial auxiliary. Furthermore, we used intonation and discourse context to diagnose 24 of the 37 auxiliary-initial utterances as assertions and 13 as questions. The sole SAux structure is an assertion. Paddy therefore uses AuxS order for both assertions and questions. We can therefore rule out a hypothesis whereby Paddy treats auxiliaries generally as question markers. The next question is why Paddy clearly prefers auxiliary-initial word order independently of speech act, despite hearing some SAux structures.

**3. Proposal: Word order mappings are subject to the Tolerance Principle**

We propose that Paddy’s word order mapping is motivated by a high relative frequency of AuxS constructions in his input, which makes him treat AuxS word order as canonical. SAux constructions in the input are treated as lexicalized exceptions or temporarily disregarded because there are not enough of them to lead Paddy to revise his AuxS rule. We frame this proposal in Yang’s (2016) Tolerance Principle combined with his (2002, 2004) account of Variational Learning. The former explains Paddy’s initial postulation of a target-deviant rule based on frequency of types in his input while the latter predicts a later stage in which rules compete, and are ultimately specialised using both input frequency and other contextual factors (cf. Wallenberg 2016, 2019). The Tolerance Principle predicts that a non-canonical variant may prevail in a learner’s grammar if the proportional frequency of exceptions in the learner’s input fails to breach a threshold as determined by a specific formula. As the input develops, however, the threshold will shift and the learner may be forced to revise their earlier rule. The common denominator between these mechanisms is the idea that early hypotheses about form-function mappings are frequency-sensitive and are subject to revision. Since learned mappings have to deal with exceptions, a combination of these mechanisms can help account for exactly when changes in rule application occur in the course of acquisition; namely after a point where input frequency reaches a certain threshold. The formula used to determine this threshold – the Tolerance Principle itself - is defined in Yang (2018: 1) as in (9):

1. Let a rule *R* be defined over a set of *N* items. *R* is productive if and only if *e*, the number of items not supporting *R*, does not exceed *θN*:

$$e\leq  θ\_{N}=\frac{N}{lnN}$$

The principle in (9) restricts the requirement for revising learning hypotheses by determining a threshold that is relative to the number of exceptions to a rule. By extension, we expect that children will only hypothesise competing rules (cf. Kroch 1989 et seq.)—i.e. enter into a process of Variational Learning—once there is sufficient evidence that exceptions to a one-to-one mapping rule cannot be analyzed as listed lexical exceptions. This means that the child now has evidence for multiple rules applying to the same set of lexical items, and neither can be reduced to a list of lexical exceptions to the other. The child has no choice but to posit two productive rules in competition (i.e. variation) with each other, and to track their probabilities as the Variational Learning model suggests. Once we apply this proposal to Paddy’s acquisition of word order mapping, the large number of auxiliary-initial utterances in his input suggests that any and every speech act that is expressed by an utterance containing an auxiliary is initially mapped onto an AuxS construction. As long as the frequency of auxiliary types in SAux constructions in the input does not exceed the threshold calculated over all auxiliary types subject to *R*, the unmarked word order for all utterances containing an auxiliary will be AuxS. Paddy’s sensitivity to the distinction between verbs and auxiliaries is crucial here since there are plenty of SVO constructions present in Paddy’s input for him to produce adult-like assertions that do not contain an auxiliary.

Let us show how the Tolerance Principle plays out for Paddy’s production given the auxiliary types that he hears. We believe that it is valid to apply the Tolerance Principle here as auxiliaries, a closed-class of lexical items with restricted syntactic behaviour, behave sufficiently like the morphosyntactic phenomena that the Tolerance Principle has been applied to successfully in past work (see e.g. Schuler et al 2016). Moreover, usage-based accounts often point out that auxiliaries are acquired incrementally (Johnson 1973), so it seems wise to control for the possibility that the positional mapping occurs for each auxiliary individually—at least in the initial stages of acquiring word order. Consider the data in Table 2 which displays all occurrences of auxiliaries by construction for Paddy’s input and output in our audio-recorded data.

|  |  |  |
| --- | --- | --- |
|  | **Input** | **Output** |
| ***AuxS*** | ***SAux*** | ***AuxS*** | ***SAux*** |
| *auxiliary BE* | 15 | 6 | 4 | 1 |
| *can* | 27 | 20 | 5 | - |
| *copula BE* | 25 | - | 18 | - |
| *could* | 1 | - | - | - |
| *do* | 28 | 3 | 6 | - |
| *have* | 12 | - | 2 | - |
| *might* | - | 1 | - | - |
| *shall* | 8 | - | 2 | - |
| *will* | - | 1 | - | - |
| **Total** | **116** | **31** | **34**[[6]](#footnote-7) | **1** |

*Table 2: Input and output frequency of auxiliary type by construction type in Paddy’s audio recordings*

Paddy hears 9 different auxiliary types[[7]](#footnote-8), of which 6 on more than one occasion. According to the Tolerance Principle in (9), 9/ln(9) = 4.1, so Paddy should only tolerate 4 exceptions to a general AuxS rule. On the face of it this does not hold, as he hears 5 auxiliary types in SAux constructions. However, two of these are *only* heard in SAux constructions (*might* and *will*), so these are the only true lexical exceptions as far as the Tolerance Principle is concerned; Paddy can easily form a productive AuxS rule tolerating two auxiliaries which are lexically specified for SAux.[[8]](#footnote-9) How do we then treat the three auxiliaries for which Paddy hears both AuxS and SAux orders in the input (auxiliary BE, copula BE, and *do)*? These cases may well force Paddy to posit a second productive rule, SAux, in variation (i.e. competition) with the AuxS rule, beginning a Variational Learning process. This analysis invites the question: isn’t the existence of this second productive rule contradicted by *copula* BE*, could, have,* and *shall*, which never appear with SAux in the input? The answer is no, because the competing SAux rule can also tolerate 4.1 exceptions out of 9 auxiliaries by the Tolerance Principle. While Paddy’s output is not sufficient to confirm the whole of this combination Tolerance Principle + Variational Learning analysis, his production of auxiliary BE in both orders is consistent with his having two rules in variation for a set of auxiliaries which includes auxiliary BE but does not include copula BE. (Note: the latter is covered by the productive AuxS rule, and listed as an exception to the competing SAux rule.) We include this example, the only SAux example in Paddy's output, in (10).

1. Mother: why did you have to wear a helmet? Where were you? In the c-
 Paddy: coal mine. Daddy **was** got a green one on (2;7)

Moreover, two more SAux utterances can be found in our written diary data[[9]](#footnote-10) with *do* (11) and *can* (12)*.*

1. [Context: Paddy has his shoes on, Mother is in bed without her shoes on*.*]
 Mother: have you got your new shoes on?
 Paddy: yes, like you didn't! (2;4)
2. [Context: P is pointing at a large doormat outside the open door of a stately home]
 Paddy: Can you [=Paddy] take your shoes off here
 Paddy: [sits down on doormat] Now you can! (2;5)

These numbers are very small, but it is notable that Paddy only uses SAux order with auxiliaries for which the input provided evidence of competing rules. This is consistent with the idea that Paddy has posited two productive rules for auxiliary syntax each with a small list of lexical exceptions; auxiliary BE, *can* and *do*, however, are not exceptions to either rule, and so can appear in both orders.

Naima and Eve, on the other hand, cannot and do not associate different word orders with different auxiliary types. Naima hears 13 different auxiliary types, of which 12 in SAux constructions and 11 in AuxS constructions.[[10]](#footnote-11) Eve also hears 13 auxiliary types, of which 12 in AuxS constructions and 11 in SAux constructions.[[11]](#footnote-12) As the Tolerance Principle places the threshold for exceptions in this case at 13/ln(13) = 5.07, they cannot form any kind of auxiliary position rule for any auxiliary from this data alone. Instead, they must associate different word orders with some other feature or semantic meaning.

To recap then, Paddy will use a combination of Variational Learning and recursive application of the Tolerance Principle to postulate something like the following:

1. Hypothesis A: There is only 1 productive rule, that all auxiliaries appear only in sentence-

 initial position.
Evidence for A: 9/ln(9) = 4.1, 5 exceptions to this rule (Application of Tolerance Principle)
Evidence against A: *auxiliary BE, do* and *can* appear in multiple syntactic configurations.
**Hypothesis A is falsified.**

1. Hypothesis B:
 i.There is a productive AuxS rule (by default applying to all auxiliaries), with 2 listed
 lexical exceptions: *might* and *will.*
 ii. There is a productive SAux rule (by default applying to all auxiliaries), with 4 listed
 lexical exceptions: *copula* BE*, could, have,* and *shall*,

 iii. The two productive rules are in variation (competition) with each other, and their
 probabilities of use need to be tracked (as in Yang 2000, 2002).

 Evidence for Bi: 9/ln(9) = 4.1, 3 exceptions to this rule (Application of Tolerance Principle)

 Evidence for Bii: 9/ln(9) = 4.1, 4 exceptions to this rule (Application of Tolerance Principle)

Evidence for Biii: *auxiliary BE, can,* and *do* all appear in both AuxS and SAux orders, so

 there must be competing rules. (Application of Variational Learning) **Hypothesis B is supported.**

**4. Independent motivation and implications for other mapping contexts**

Paddy’s postulation of an AuxS word order irrespective of speech act is only a temporary phenomenon. From around 2;9 onwards, after recordings have ceased, Paddy’s word order mappings change to correspond with adult canonical conventions. Our observation that the auxiliary-initial assertions in his early output correlate with an asymmetry of SAux and AuxS utterances in his early input led to our proposal that word order acquisition and mapping are sensitive to changes in input frequency: until the number of a given type of auxiliary reaches a threshold, as predicted by the Tolerance Principle, the child may postulate a non-adult-like rule based on this evidence and treat exceptions to this rule as lexicalized. Once this threshold is reached, the child revises their postulated rule. This may then prompt the child to consider two (or more) productive rules in variation, applying across lexical types. This competition between rules across lexical items can be resolved through Variational Learning, which may lead to one rule disappearing, or to the two rules specialising for non-overlapping syntactic contexts (Wallenberg 2016, 2019); this latter is arguably what happens as SAux and AuxS become restricted to non-overlapping contexts of illocutionary force (and, we claim, what Paddy will do after the temporal snapshot we consider here). Moreover, there is independent evidence that the combination of learning mechanisms proposed here apply elsewhere, too.

An independent piece of evidence that supports a variation-based learning of word order mapping is found in diachronic research. Mapping questions onto constructions with sentence-initial verbal elements is a key property of Germanic languages. Modern-day English has auxiliary-initial word order in canonical polar questions; modern-day German has a verb-initial word order in the same context. It can be assumed that this word order difference has roots in two different transformations: English moves the auxiliary to the sentence-initial position (known as T-to-C movement) while German first moves the verb into the auxiliary position (V-to-T movement) and then progresses to T-to-C movement. Earlier stages of English also had both types of movement as is exemplified in (15-16) from Roberts (1985).

1. [CP Must1 [TP they ***t*1** [VP leave?]]] (Roberts 1985: 21)
2. [CP Wilt [TP thow ***t*1** [VP ***t*1** ony thinge with hym?]]]

  *Do you want him for anything*? (1470-85 Malory Morte d'Arthure III, iii, 102, V 559)

This diachronic development very much resembles what is attested for Faroese and Mainland Scandinavian, which also lost V-to-T movement. Heycock & Wallenberg (2013) show on the basis of a quantitative corpus analysis that the modern Mainland Scandinavian grammar, which lacks V-to-T movement, will win out over time once it comes into competition with the older V-to-T grammar in languages like Faroese and Mainland Scandinavian. In such languages, a certain frequency of V-to-C verb movement in certain embedded clauses masks instances of the V-to-T grammar, restricting its ability to be tracked by a Variational Learner. Faroese is of particular interest here because the change is not yet complete. The Old Swedish sentences in (17-18) from Heycock & Wallenberg (2013:129) perfectly capture the paradigm in flux with *not* marking the left edge of the verb phrase: (17) has V-to-T movement, (18) has not. Variational Learning predicts that the grammar with the fewest ambiguous structures will emerge as the winner, which in this case is the grammar without T-to-C, as both main and embedded clause order can be captured this way.

1. æn min guþ brytar **eigh** niþar þin guþ.

 *if my god breaks* ***not*** *down your god*

 ‘if my god doesn’t break down your god’ (Old Swedish, Codex Bureanus, c. 1350)

1. om thenne her ærende **ekke** faa en snar ænda.

 *if these errands* ***not*** *get an immediate end*

 ‘if these errands do not come to an immediate end’ (Old Swedish, c. 1470–1490)

This process bears some resemblance to the problem Paddy faces at the early stage of determining the word order of his target variety, as he has more support for an AuxS word order. Yet, the Tolerance Principle threshold for an AuxS order is breached, unless he posits two productive rules in variation. As two auxiliaries, *might* and *will,* only appear in sentence-medial position, he treats them as lexicalized exceptions. However, as Paddy’s linguistic experience expands and he encounters a sufficient amount of different auxiliary types in each possible position, there will be no way for him to treat all SAux structures as exceptions to a single productive rule. We suggest he therefore posits a competing SAux rule with its own lexical exceptions applying across auxiliary types. However, in time he will resolve this competition, and he will need to decide whether his AuxS rule is still productive, in particular in cases of ambiguous data. By ambiguous data, we point to cases of auxiliary-medial questions in child input such as declarative questions (e.g. Estiarribia 2010; Kania 2016). Indeed, we know from recent investigations of the relation between clause type and speech act in the input that less than half of the questions heard by children aged 1 to 3 have AuxS word order (Zaitsu et al. 2020).[[12]](#footnote-13) More importantly, only 12% of these questions unambiguously match the biased use conventions of rising declaratives (see Gunlogson 2008; Rudin 2018; Heim 2019). Hence, there is ample opportunity for children to start out with an unconventional mapping of SAux word order to questions. Eventually, however, the number of auxiliary-medial assertions becomes significantly larger than that of the auxiliary-medial questions. What once was treated an lexicalised exception is now the canonical word order for assertions.

What we claim here, then, is that Paddy’s auxiliary-initial assertions may well be exceptional with respect to what is recorded in the acquisition literature, but they are exceptional not because Paddy is himself exceptional. Rather we claim the conditions for recording them were (accidentally) met. These conditions are a combination of early multiword utterances and a strong asymmetry of SAux and AuxS word order (crucially towards AuxS order) in the input during the initial stages of acquisition. The fact that Paddy transitioned to auxiliary-medial assertions at a later point as well as the diachronic support for a competition between grammars in Scandinavian based on differences in word order requirements, both of which are correctly predicted by Variational Learning, lead us to believe that our proposal identifies a feature of language acquisition that is not restricted to the morphosyntactic domain. The mapping of speech acts onto their canonical clause types is underpinned by the same combination of frequency-based learning mechanisms as any form-function mapping that needs to deal with (lexicalized) exceptions.

**5. Conclusion and a call for further research**

Based on a small sample of caregiver-infant exchanges involving a cognitively typical early talking acquirer of British English, we have argued that word order mappings are learned and sensitive to input frequency. The data we presented suggests that English-acquiring children may postulate a non-canonical default word order for utterances containing auxiliaries if they have more evidence in the input for AuxS than SAux according to an input-relative threshold. However, we will only find evidence for this non-adult-like reasoning if they start *producing* sentences containing auxiliaries during this period, which seems to occur very early in the acquisition process. The fact that this mapping changes once a certain threshold is exceeded (i.e., as the child’s input becomes more diverse) further suggests that possible grammars only compete in a Variational Learning sense if a (set of) productive rule(s) cannot be determined according to the frequency of morphosyntactic forms in the input alone. A close examination of Paddy’s input revealed that, given the number of auxiliary types he heard, Paddy could first postulate a general AuxS word order rule, the exceptions to which were fewer than the threshold predicted by the Tolerance Principle (Yang 2016; 2018), followed by a secondary SAux rule for a certain proportion of these types. What we propose here, then, is that the combination of Variational Learning and the Tolerance Principle can explain initial mappings both in syntactic development and syntactic change. To see this full process in action with respect to word order acquisition, we must identify linguistically precocious children who vary with respect to the relative frequency of AuxS and SAux structures in their input. We are in the process of recording and transcribing the input and production of two more such children, but we call upon other researchers to engage in searching for similar paradigms and learners that can help expand our understanding of early rule formation, Variational Learning and its relation to hypothesis revision in language acquisition.

**Acknowledgements**

We are grateful to Robyn Orfitelli for her work on an earlier version of this project, presented at the Linguistics Association of Great Britain annual meeting in September 2019, and to Tom Roeper for his ever helpful and fruitful discussion. We also thank the attendees of WCCFL 2021 for their time and feedback, especially Heidi Harley, Max Papillon, and Charles Yang. Finally, a huge thank you to Paddy and his family, especially his mother, who made the audio recordings, for their time and generosity.

**References**

Brown, Roger. 1973. *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Brown, Roger, & Bellugi, Ursula. 1964. Three processes in the child's acquisition of syntax. *Harvard Educational Review*, 34(2), 133-151.

Charney, Rosalind. 1980. Speech roles and the development of personal pronouns. *Journal of Child Language* 7, 509-528.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry*, 36(1):1–22.

Dale, Philip S. & Catherine Crain-Thoreson. 1993. Pronoun reversals: who, when and why? *Journal of Child Language* 20(3), 573-589. DOI: doi:10.1017/S0305000900008485

Estigarribia, Bruno. 2010. Facilitation by variation: right‐to‐left learning of English yes/no questions. *Cognitive Science* 34(1), 68-93.

Evans, Karen, & Demuth, Katherine. 2012. Individual Differences in Pronoun Reversal: Evidence from two Longitudinal Case Studies. *Journal of Child Language* 39: 162-191.

Gunlogson, Christine. 2003. *True to form: Rising and falling declaratives and questions in English*. London: Routledge.

Heim, Johannes. 2019. *Commitment & Engagement: The role of intonation in deriving speech acts*. Doctoral dissertation, University of British Columbia.

Heycock, Caroline & Joel C. Wallenberg. 2013. How variational acquisition drives syntactic change. *The Journal of Comparative Germanic Linguistics* 16(2-3), 127-157.

Hirsh-Pasek, Kathy & Roberta M. Golinkoff. 1996. The intermodal preferential looking paradigm: A window onto emerging language comprehension. In Dana McDaniel, Cecile McKee & Helen S. Cairns (eds), *Methods for assessing children’s syntax* (pp. 105–124). Cambridge, MA: MIT Press.

Johnson, Carolyn. E. 1983. The development of children’s interrogatives: From formulas to rules. *Papers and Reports on Child Language Development* 22, 108–115.

Kania, Ursula. 2016. *The acquisition and use of yes-no questions in English: A corpus-study from a usage-based perspective* (Vol. 36). Tübingen: Narr Francke Attempto Verlag.

Kroch, Anthony S. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1(3), 199-244.

MacWhinney, Brian. 2000. *The CHILDES Project: Tools for analyzing talk. Third Edition.* Mahwah, NJ: Lawrence Erlbaum Associates.

Roberts, Ian. 1985. Agreement parameters and the development of English modal auxiliaries. *Natural Language and Linguistic Theory* 3: 21–58.

Rudin, Deniz. 2018. *Rising above commitment.* Doctoral dissertation, University of California Santa Cruz.

Sano, Tetsuya & Nina Hyams. 1994. Agreement, finiteness, and the development of null arguments. In *Proceedings of NELS.* Vol. 24. Amherst, MA: GLSA.

Schuler, Kathryn D., Charles Yang & Elissa L. Newport. 2016. Testing the Tolerance Principle: Children form productive rules when it is computationally more efficient to do so. In Anna Papafragou, Daniel J. Grodner, Daniel Mirman & John Trueswell (eds) *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 2321–2326).Austin, TX: Cognitive Science Society.

Schütze, Carsten, & Kenneth Wexler. 1996. Subject case licensing and English root infinitives. In *Proceedings of the 20th annual BU Conference on Lang. Development* (Vol. 2, pp. 670-681). Somerville, MA: Cascadilla Press.

Stromswold, Karin. J. 1990. *Learnability and the acquisition of auxiliaries.* Doctoral dissertation, Massachusetts Institute of Technology.

Theakston, Anna L., Elena V. M. Lieven, Julian M. Pine & Caroline F. Rowland. 2001. The role of performance limitations in the acquisition of verb-argument structure: an alternative account. *Journal of Child Language* 28, 127-152.

Wallenberg, Joel C. 2016. Extraposition is disappearing. *Language* 92(4), e237-e256.

Wallenberg, Joel C. 2019. A variational theory of specialization in acquisition and diachrony. *The Determinants of Diachronic Stability* 254, 245.

Yang, Charles. 2000. Internal and external forces in language change. *Language variation and change*, 12(3), 231-250.

Yang, Charles. 2002. *Knowledge and learning in natural language*. Oxford: Oxford University Press.

Yang, Charles. 2004. Universal Grammar, statistics or both?. *Trends in Cognitive Sciences* 8(10), 451-456.

Yang, Charles. 2016. *The price of linguistic productivity: How children learn to break rules of language.* Cambridge, MA.: MIT Press.

Yang, Charles. 2018. *A User’s Guide to the Tolerance Principle*. lingbuzz/004146

Zaitsu, Anissa, Jad Wehbe, Valentine Hacquard & Jeff Lidz. 2020. Clause types and speech acts in speech to children. Presentation at Experiments in Linguistic Meaning (ELM) 1.

Zane, Emily, Sudha Arunachalam & Rhiannon Luyster. 2020. Caregiver-Reported Pronominal Errors Made by Children with and without Autism Spectrum Disorder. In Danielle Dionne and Lee-Ann Vidal Covas (eds), *Proceedings of the 45th annual Boston University Conference on Language Development* (pp. 845-859). Somerville, MA: Cascadilla Press.

1. Rebecca Woods, Newcastle University: rebecca.woods@ncl.ac.uk. Johannes Heim, Newcastle University: johannes.heim@ncl.ac.uk. Joel Wallenberg, Newcastle University: joel.wallenberg@ncl.ac.uk. [↑](#footnote-ref-2)
2. Reviewers and commenters have suggested that Paddy’s tendency to reverse pronouns, especially the second person singular, are indicative of neurodivergence. We strongly refute this on the basis of (a) Paddy’s individual history and evidence of pronoun reversal in other neurotypical children in both corpus studies, particularly of linguistically precocious children (e.g. Dale and Crain-Thoreson 1993, Evans and Demuth 2012) and experimental studies (e.g. Charney 1980). The fact that pronoun reversal is more common in children with autism than in typically developing children (e.g. Zane et al., 2021) is not relevant here. [↑](#footnote-ref-3)
3. This also holds for the 12-child Manchester corpus (Theakston et al 2001). [↑](#footnote-ref-4)
4. Note that Eve and Naima are both acquiring American English as spoken on the East Coast of the United States, though we do not believe this to be pertinent to the arguments made herein. [↑](#footnote-ref-5)
5. For the purpose of quantitative comparison, we exclude written diary data, as we principally only have this for Paddy and not for his caregivers. [↑](#footnote-ref-6)
6. Note that 3 auxiliaries are missing from this table: 3 instances of BE could not be definitively coded as either copula or auxiliary BE as they were contained within partial utterances. [↑](#footnote-ref-7)
7. We assert that Paddy does, in fact, treat the items in table 2 as a single part of speech, namely auxiliaries, for a number of reasons: they are differentiated from lexical verbs (which Paddy never inverts), those which can bear other morphemes (e.g. past tense morphology) do so, and they always appear with a nominative-marked subject. [↑](#footnote-ref-8)
8. See Van Valin (2002: 167) for a similar proposal, where some modals are treated differently from BE/HAVE/DO by the English-acquiring child in the earliest stages of auxiliary acquisition. [↑](#footnote-ref-9)
9. There are also two examples in the form “There pronoun BE” (e.g. “There you are!”), but we suggest that these are frozen forms familiar from and used in e.g. peekaboo games rather than truly novel, generated utterances. [↑](#footnote-ref-10)
10. SAux: auxiliary BE*, do, have, can, would, could, might, must, should, shall, may, will.*
AuxS: copula BE*, can, should, do,* auxiliary BE*, have, would, could, shall, may, will.*  [↑](#footnote-ref-11)
11. SAux: *can, do, will, have,* auxiliary BE*, might, may, should, would, could, must.*
AuxS: *would,* auxiliary BE*,* copula BE*, will, can, do, have, may, should, could, shall.* [↑](#footnote-ref-12)
12. Interestingly, only one SAux construction in Paddy’s audio-recorded input can be construed as a question, in this case a quiz-style question where the queried constituent is left entirely unexpressed: “and then in the back you can make?” This lack of SAux questions in Paddy’s input will again feed—or at least fail to curb—his early postulation of a general AuxS rule; this example is also consistent with the later SAux rule that we claim Paddy then postulates. [↑](#footnote-ref-13)