The 'human-like' learnability of the *wh*- and coordination island constraints by artificial neural networks in Dutch

Michelle Suijkerbuijk, michelle.suijkerbuijk@ru.nl, Radboud University
Stefan Frank, stefan.frank@ru.nl, Radboud University
Peter de Swart, peter.deswart@ru.nl, Radboud University

The current study investigated whether an artificial neural network (ANN) can learn the *wh*- and coordination island constraints in Dutch, and directly compared its performance to that of native speakers. First, speakers' sensitivity to these constraints was assessed with an acceptability judgement task, and subsequently the same sentences were presented to an ANN. While the Dutch speakers seem to show a sensitivity to *wh*- and coordination island violations, the ANN does not seem able to similarly learn to recognize these gap-resistant structures.

1 Introduction

Nearly all children acquire the syntax of their first language with ease, but how is that actually possible? While some have argued that humans need innate knowledge of language (Chomsky, 1971), recent research suggests that artificial neural networks (ANNs) can induce human-like grammatical knowledge without having linguistic abilities built in (Linzen & Baroni, 2021). However, this recent research almost all (1) use English as the input language, and (2) do not directly compare the ANN's performance to that of actual native speakers, which makes it impossible to state that the ANN has acquired knowledge that is 'human-like'. Therefore, we investigated whether an ANN can learn syntactic constraints **in Dutch**, and **directly compared the ANN's performance to that of Dutch native speakers**. Specifically, we examined the learnability of syntactic island constraints.

2 Theoretical background

Syntactic island constraints are conditions on non-local dependency relations, prohibiting movement out of syntactic islands (Liu et al., 2022), such as *wh*-phrases or coordinations. These constraints have played an important role in the development of syntactic theories, and their predominant analysis in these theories relies heavily on the assumed innate language ability (Pearl & Sprouse, 2013). An example of a *wh*- and a coordination island violation can be found in (1), taken from Liu et al. (2022, p. 497).

(1) a. *What did you wonder [$_{\text{wh-island}}$ whether John bought <*what*>]?
a. *What did John buy [$_{\text{coordination island}}$ a shirt and <*what*>]?

With regard to research on syntactic island constraints, either with ANNs or native speakers, Dutch is an underrepresented language. Previous studies with ANNs on island constraints have mainly been performed in English and show mixed results; some islands, such as *wh*- and coordination islands, are learned successfully, but others only partially or not at all (e.g., Wilcox et al., 2021). As it would be interesting to see whether these successfully learned constraints in English can also be learned in another language, typologically similar but also different from English, and underrepresented in the current state of affairs, the current study investigated whether these two island types could also be successfully learned by ANNs in Dutch. First, however, experimental data of native Dutch speakers on these island types had to be gathered as not much is known about whether these island constraints exist in Dutch, and if so, to what extent its speakers are sensitive to them. Beljon et al. (2021) is one of the few, if not the only, study that empirically investigated Dutch native speakers' sensitivity to islands, specifically to *wh*-islands, and showed that they are strongly sensitive to the *wh*-island constraint. For the coordination island constraint, it has been argued that it is never possible, in any language, to extract an element out of (part of) a conjunct, although there is almost no

experimental evidence to support or oppose that claim (Liu et al., 2022). Due to this (near) absence of experimental evidence for *wh*- and coordination islands in Dutch, the current study gathered human experimental data to first establish whether the *wh*- and coordination island constraints exist in Dutch and if so, to what extent native speakers are sensitive to them. Next, to find out more about the cross-linguistic abilities of the ANNs, it was investigated whether a network demonstrates a human-like sensitivity to the island constraints in Dutch.

3 Methodology

3.1 Experimental design and materials

The interaction design used in the current study (based on Wilcox et al., 2021) builds on two predictions assumed to be made by the grammar: (1) gaps require fillers, and (2) fillers require gaps. Consequently, the independent variables PRESENCE OF GAP and PRESENCE OF FILLER were included in the design, resulting in four conditions, shown in a regular declarative sentence in Table 1.

Table 1

Conditions of PRESENCE OF **FILLER** and PRESENCE OF GAP in a regular declarative sentence.

Gap?	Filler?	Example sentence
No	No	*Ik weet **dat** jij zag dat de bakker koekjes maakte in de bakkerij.* I know that you saw that the baker cookies made in the bakery
No	Yes	**Ik weet **wat** jij zag dat de bakker koekjes maakte in de bakkerij.* I know what you saw that the baker cookies made in de bakery 'I know (*w/t)hat you saw that the baker made cookies in the bakery.'
Yes	No	**Ik weet **dat** jij zag dat de bakker ___ maakte in de bakkerij.* I know that you saw that the baker GAP made in the bakery
Yes	Yes	*Ik weet **wat** jij zag dat de bakker ___ maakte in de bakkerij.* I know what you saw that the baker GAP made in the bakery 'I know (w/*t)hat you saw that the baker made in the bakery.'

Furthermore, the independent variables ISLAND (non-island vs. island) and STRUCTURE (*wh*-island vs coordination island) were added to the design to compare non-islands to two island types. An example of the [-gap, -filler] condition within a *wh*-island can be found in (2).

(2) *Ik weet dat jij je afvraagt of de bakker koekjes maakte in de bakkerij.*
I know that you REFL[1] wonder whether the baker cookies made in the bakery
'I know that you wonder whether the baker made cookies in the bakery.'

Additionally, as it has been suggested that native speakers and ANNs simply cannot thread information through the syntactically complex islands (Chowdhury & Zamparelli, 2018; Keshev & Meltzer-Asscher, 2019), control items were added to the experiment, in which participants had to maintain expectations for gendered pronouns either through a *wh*-/coordination island or a non-island configuration. An example of a *wh*-island control item can be found in (3).

(3) *Ik weet dat de meester zich afvraagt of*
I know that the male.teacher REFL wonders whether
de leerlingen (zijn/?haar) uitleg begrijpen.
the students his/her explanation understand
'I know that the teacher wonders whether the students understand (his/?her) explanation.'

[1] REFL = reflexive pronoun

3.2 Participants, procedure and hypotheses

88 native Dutch speakers (M_{age} = 19.61, range 17-33, SD_{age} = 2.04) judged 160 sentences (64 experimental items, 32 control items, 64 fillers) on their acceptability in Dutch on a scale from 1 (*Erg slecht* 'very bad') to 7 (*Erg goed* 'very good'). Moreover, a Long Short-Term Memory (LSTM) network, trained on 12 million sentences extracted from the Dutch *Corpora Of the Web* (NLCOW14, Schäfer, 2015), assigned surprisal values to the same experimental and control items, indicating the extent to which a word was unexpected by the network.

Before the creation of this experiment, the following hypotheses were made. First, within regular filler-gap dependencies, gaps should be less acceptable and more surprising when no *wh*-filler is present. Within island configurations, however, it should always be ungrammatical to create a gap within an island configuration, regardless of the presence of a filler. Second, filled argument positions should be less acceptable and more surprising when a *wh*-filler is present both within regular filler-gap dependencies and within island configurations; encountering a *wh*-filler should always give rise to the expectation of a gap, and not encountering this gap anywhere in the sentence should make the sentence less acceptable and more surprising.

4 Results

4.1 Acceptability judgement task

To start, the control items showed that Dutch native speakers can maintain gender expectancies through island configurations, as the gender matches were rated as more acceptable than the mismatches. This shows that they are capable of threading information through these complex structures.

The standardized acceptability judgements, illustrated in Figure 1A, were analysed in an LMER model, which revealed a significant interaction between PRESENCE OF FILLER, PRESENCE OF GAP and ISLAND (β = −.04, SE = .02, 95% CI of β = [−.08, −.00], p = .030). First, for regular filler-gap dependencies, the figure shows that gaps are judged as less acceptable with no *wh*-filler present in the sentence, and that filled argument positions are perceived as less acceptable when there is a *wh*-filler in the sentence. Second, for island configurations, it can be seen that the presence of a *wh*-filler decreases the acceptability ratings of both sentences with and without gaps within the islands.

4.2 Long Short-Term Memory network

The surprisal values measured on the immediate post-gap verb, illustrated in Figure 1B, were analysed in an LMER model, which also revealed a significant interaction between PRESENCE OF FILLER, PRESENCE OF GAP and ISLAND (β = 1.79, SE = .66, 95% CI of β = [.50, 3.10], p = .007). The figure shows exactly the same pattern for both regular filler-gap dependencies and island configurations; gaps are judged as more surprising with no *wh*-filler present in the sentence, but surprisal values do not seem affected by the presence of a *wh*-filler in sentences with filled argument positions. The interaction effect found thus seems to be solely driven by the difference in effect size as opposed to effect direction. Also important to note here is that the control items with gender expectations remained inconclusive, and can thus not provide the control that was aimed for.

5 Discussion and theoretical implications

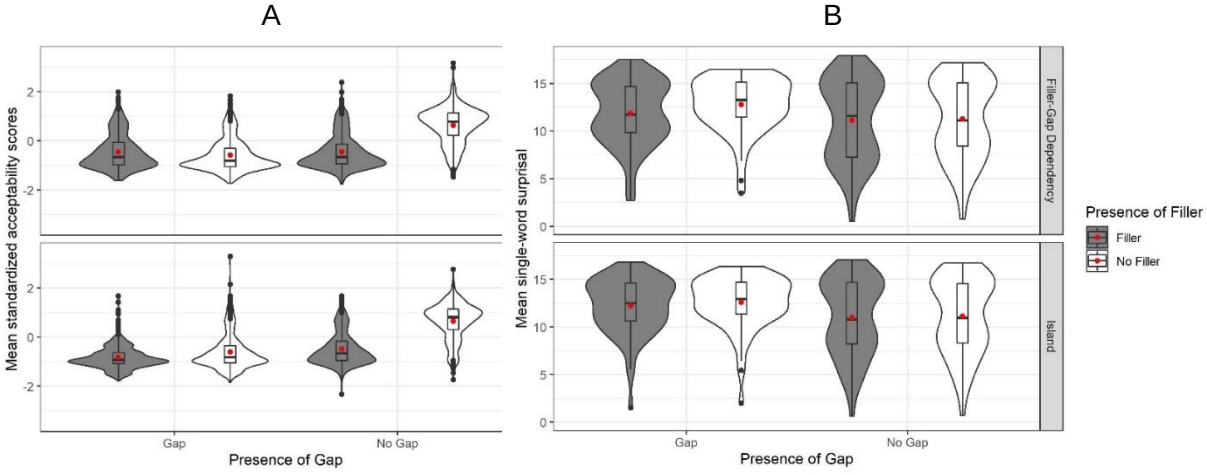
While the human native speakers of Dutch show sensitivity to *wh*- and coordination island violations, the LSTM network did not learn to recognize these gap-resistant structures. This could be due to various reasons either specific to this research project, such as the complexity of the items used, the analysis of the surprisal values and the training of the neural network architecture, or due to a more general reason, such as the structural properties of Dutch or the absence of an innate language ability in the network. While future research should first resolve the research-specific issues, it is still interesting to discuss the latter explanation.

While children can already recognize syntactic islands at the age of four, the network was unable do the same with the current training data covering a lot more than four years of a human's life (Wilcox et al., 2021). This suggests that the information in the training data was not enough for the network to learn about syntactic island constraints, and that children might thus use something else than just external input to learn the syntactic island constraints not available to the network (e.g. internal language knowledge/abilities).

While this research thus provides relevant new insights for the debate about language acquisition, it also provides relevant new knowledge to the field of experimental syntax. The current results show that the *wh*- and coordinate structure island constraints exist in Dutch, which strengthens the results found by Beljon et al. (2021) and experimentally supports the theoretical claim that "it does not seem possible to extract one or more full conjuncts" (Liu et al., 2022, p. 503), at least for Dutch.

Figure 1

(A) Violin/boxplot of standardized acceptability ratings; and (B) mean single-word surprisal values; as a function of PRESENCE OF GAP, PRESENCE OF FILLER, and ISLAND.



References

Beljon, M., Joosen, D., Koeneman, O., Ploum, B., Sommer, N., de Swart, P., & Wilms, V. (2021). The effect of filler complexity and context on the acceptability of the wh-island violations in Dutch. In M. Dingemanse, E. Van Lier, & J. Vogels (Eds.), *Linguistics in the Netherlands* (Vol. 38, pp. 4–20). John Benjamins Publishing Company.

Chomsky, N. (1971). *Problems of Knowledge and Freedom*. Fontana.

Chowdhury, S.A., & Zamparelli, R. (2018). RNN simulations of grammaticality judgments on long-distance dependencies. In *Proceedings of the 27th International Conference on Computational Linguistics* (pp. 133–144). Association for Computational Linguistics.

Keshev, M., & Meltzer-Asscher, A. (2018). A processing-based account of subliminal wh-island effects. *Natural Language & Linguistic Theory*, *37*(2), 621–657.

Linzen, T., & Baroni, M. (2021). Syntactic structure from deep learning. *Annual Review of Linguistics*, *7*(1), 195–212.

Liu, Y., Winckel, E., Abeillé, A., Hemforth, B., & Gibson, E. (2022). Structural, functional, and processing perspectives on linguistic island effects. *Annual Reviews*, *8*, 495–525.

Pearl, L., & Sprouse, J. (2013). Computational models of acquisition for islands. *Experimental Syntax and Island Effects*, 109–131.

Schäfer, R. (2015). Processing and querying large web corpora with the COW14 architecture. In *Proceedings of the 3rd Workshop on Challenges in the Management of Large Corpora (CMLC-3)* (pp. 28–34). Institut für Deutsche Sprache.

Wilcox, E.G., Vani, P., & Levy, R.P. (2021). A Targeted Assessment of Incremental Processing in Neural Language Models and Humans. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing* (pp. 939-952). Association for Computational Linguistics.