Comments on pre-reading quiz 6

1. In your own words, what is the structural difference between a recursively consistent and recursively inconsistent grammar?

A recursively consistent grammar has uniform head direction (head initial or head final) in phrases that can recursively contain each other. For example, an NP can contain a PP and a PP can contain an NP. NPs can be head-initial (Noun then other stuff) or head-final (stff then Noun); likewise, PPs can be head-initial or head-final. A recursively consistent language would have both NPs and PPs being head-initial or head-final; a recursively inconsistent language might have a head-initial NP and a head-final PP, or the reverse.

2. Which of the following statements captures the simulation results from Christiansen & Devlin's neural network model?

☑ Grammars exhibiting recursive inconsistency lead to higher Mean Squared Error after a fixed period of training.

◯ Grammars exhibiting recursive inconsistency lead to lower Mean Squared Error after a fixed period of training.

Figure 5 shows that more inconsistent languages have higher MSE. Remember that means higher MSE means that the network is not very good at predicting upcoming syntactic categories based on the sequence so far. So this result shows that the networks tend to do less well at learning inconsistent languages (given a fixed training period - I guess they might eventually catch up if you trained them enough).

3. Which of the following statements captures Christiansen & Devlin's analysis of cross-linguistic variation?

◯ Grammars exhibiting recursive inconsistency are relatively common

☑ Grammars exhibiting recursive inconsistency are relatively rare

They split their analysis by PP and PossP, but in both cases it looks like the consistent languages make up the vast majority of the languages in their sample - 94% for the adpositions, 82% for the possessives. In other words, recursively inconsistent languages seem to be quite rare.

4. Why do you think Christiansen & Devlin use an SRN to test the learnability of different grammars?

☐ SRNs have language-specific knowledge built in.

☑ SRNs do not have language-specific knowledge built in.

☑ Running neural network simulations is easier than running experiments with humans.

☑ Unlike basic feedforward networks, SRNs are capable of sequence learning

☐ Like basic feedforward networks, SRNs are capable of sequence learning

I think the crucial one for their rhetoric is that SRNs do **not** have language-specific knowledge built in. You might want to argue that the scarcity of recursively-inconsistent languages is somehow a reflection of a specialised human capacity for language learning - a language instinct that says "don't learn a recursively inconsistent language" or something like that. However, if an SRN has the same bias against recursive inconsistency, that shows that we don't need a special explanation for that behaviour in language learning in humans, since SRNs can learn any old sequential pattern you like; furthermore, it might suggest (more of a stretch here!) that humans are applying some kind of domain-general sequence learning to learn language.

But it's also true that running simulations is easier than running experiments (or once you have set up your model it's easier!). Later in the course we'll see an experimental result with humans similar in flavour to Christiansen & Devlin's SRN results.

And finally, crucially, SRNs are capable of learning sequences, thanks to their context layer, which enables them to 'remember' elements of a sequence prior to the current one.

5. In your own words, explain why learnability of a language might be correlated (inversely) with the frequency of that language type in the world's languages.

We are going to spend more time thinking about this - in many ways this is the crucial question that this course addresses, so I will be interested to read what you think at this very early stage. At this point I am expecting you to say something like "well, languages have to be learned, so maybe the ones that are hard to learn die out or are likely to be replaced by ones that are easier to learn", which I think is roughly what Christiansen & Devlin have in mind.