Prosody in sentences: Some comments

Michael Wagner (McGill University)

LabPhon 19. Hanyang University, Seoul. June 27, 2024

1. Nele Ots: Cross-linguistic survey of intonation planning: a cognitive approach
2. Guo et al.: The Garden Path Leading to Intonational Phonology
3. Ferreira: Prosody, Syntax, and Conversational Language

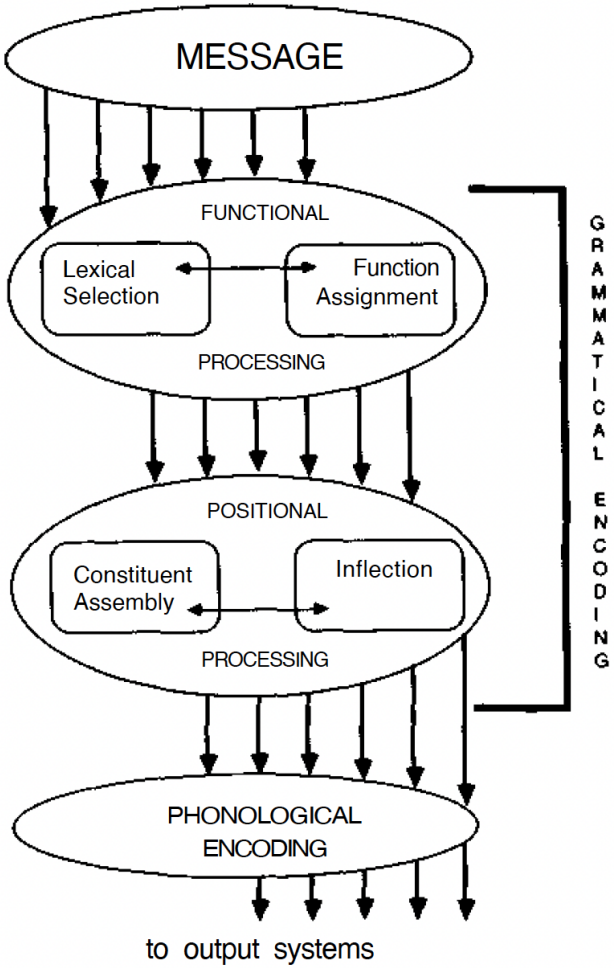


FIG. 1 An overview of language production processes.

(Bock and Levelt, 1994)

- A message is first mapped to functional structure (e.g. chose active over passive), then planned from left to right (positional processing)
- Planning constituents is incremental [1]planning phonological detail even much more so!):
  - In English SVO, SV is planned before speech onset, but not O (e.g. Lindsley, 1975)
  - Initial pitch tends to scale higher with longer upcoming material, but only length of initial constituents matter Fuchs et al. (2013)
  - Constituents with more 'available' referents tend to be ordered first (by choosing passive over active; or by reordering (e.g., 'heavy-shift', 'scrambling', etc.)

[1](

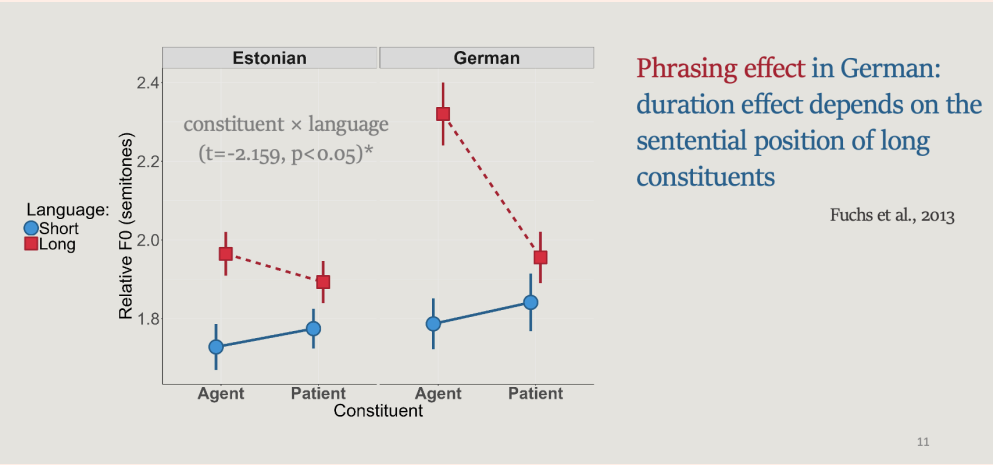
Do speakers of flexible-word order languages plan ahead more or less?
We might expect less:

- Rigid word order preempts ordering available information early
- Flexible word order enables it, facilitating incremental planning
- E.g. V. Ferreira (1996), looking at flexible vs. inflexible ditransitives:

  "When constructing a sentence allowed a syntactic choice, speakers generally constructed that utterance with fewer errors and more quickly."

- **But actually:** It takes *longer* to initiate speaking in Russian compared to English (Myachykov et al., 2013)—although Russian is more flexible (and norming showed speakers make use of this greater flexibility in the very stimuli used)
- **What about Estonian and German?**

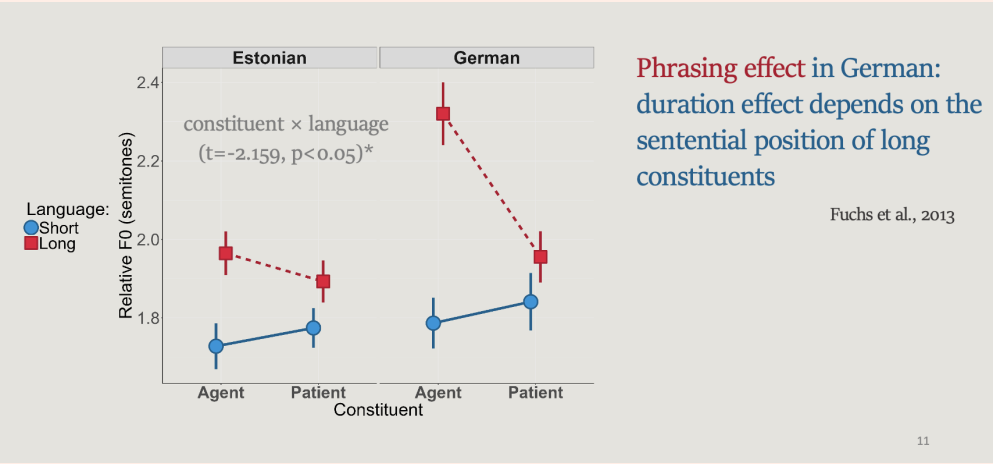
Ots: Length effects on pitch scaling



"The planning increment is larger for Estonian than for German"

- The plot suggests that the interaction (constituent*language) might mostly be driven by a smaller length effect due to agents in Estonian, rather than a smaller length effect due to patient in German
- A more direct test of the hypothesis would be to look for an interaction just based on patient-length effects on F0 in the two languages (since patient-length effects are evidence for advance planning)
- The smaller length-effect of agents in Estonian might alternatively reveal that only part of the first NP was planned (Estonian more incremental?)

Ots: Effects of length on pitch

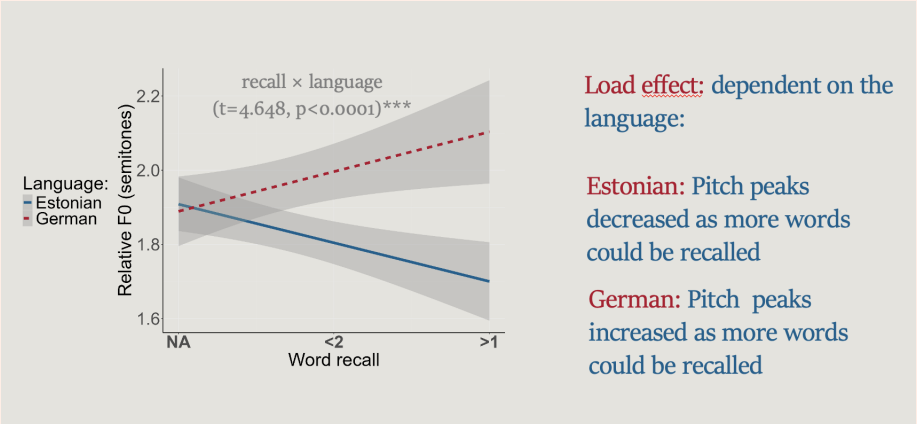


"The planning increment is larger for Estonian than for German"

Potential further tests of the hypothesis:

- How did the languages compare with respect to agent vs. patient length effects on utterance initiation time (which is what Lindsley looked at)?
- And: did pitch raising correlate with UIT within language?
- This could establish that both measures reflect the same *level* of planning: Is it advance planning of the *utterance* or advance planning of the *message* (cf. Brown-Schmidt and Konopka, 2014, i.a.)?

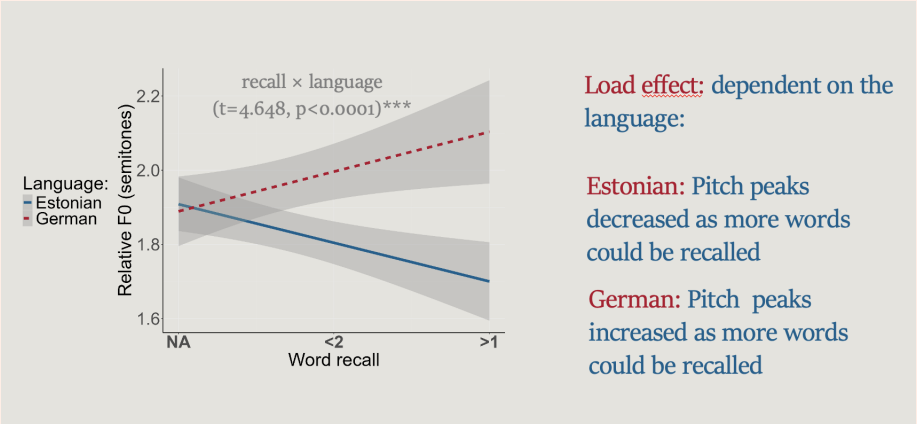
Ots: Effects of cognitive load



"F0 lowering in Estonian indicates internal load on (verbal) working memory
F0 raising in German suggests external load on working memory."

- It seems surprising to me that the same task poses different cognitive challenges depending on language
- Could it be that the reverse effects in the two languages reflect differences in the intonational structure?
- Ots, p.c., suggests that in Estonian, speakers tended to use H* accents, and in German, LH* accents—maybe cognitive load leads to undershoot?

Ots: Effects of cognitive load



"F0 lowering in Estonian indicates internal load on (verbal) working memory
F0 raising in German suggests external load on working memory."

Potential further tests of the hypothesis:

- Cognitive load shrinks the planning window (Ferreira and Swets, 2002; Swets et al., 2007; Konopka, 2012)—so it should modulate the the length effect on F_0—does it?
- Cognitive load should also modulate length-effects on UIT, and possibly word order choices (word order choice is not reported here yet)
- Testing for these effects could further strengthen the argument

A higher level conclusion:

One source of variability in sentence prosody (in the case here: pitch scaling):

The window size of production planning is variable

1. Nele Ots: Cross-linguistic survey of intonation planning: a cognitive approach

2. Guo et al.: The Garden Path Leading to Intonational Phonology

3. Ferreira: Prosody, Syntax, and Conversational Language

1

It was the humorist that ___ was leaving the scene

This could answer one of two kinds of questions, revealing a structural ambiguity:

Connected Clause Parse (CC) 2

a. Who was leaving the scene?
It was ∼[[the HUMORIST]$_F$ that ___ was leaving the scene]

Restrictive relative clause Parse (RRC) 3

b. Who called?
It was ∼[the HUMORIST ___ was leaving the SCENE]$_F$]

Guo et al:

- There's a bias toward the CC parse in (a)
- **Prosody helps resolve the ambiguity** (although both end in an RC)
- The prosodic facts support metrical theories of prominence

Why the prosodic difference? The production study made the RC material given in CC structure:

Connected Clause Parse (CC) 4

a. Who **was leaving the scene**?
It was ∼[[the HUMORIST]$_F$ that ___ **was leaving the scene**]

But not in RRC structure:

Restrictive relative clause Parse (RRC) 5

b. Who called?
It was ∼[the HUMORIST ___ **was leaving the SCENE**]$_F$

- But the same prosodic difference seems likely without context—**Why?**

Connected Clause Parse (CC) 6

a. Who **was leaving the scene**?
It was ∼[[the HUMORIST]$_F$ that ___ *was leaving the scene*]

- The relative clause encodes the question that the pivot of the cleft answers
- In the absence of a context, we seem to assume that this question is the already the salient QUD (why else was the utterance made?)
- But this is actually just a tendency (Williams, 1997; Wagner, 2012; Hedberg, 2013; Büring, 2016, i.a.)—the RC in clefts can also encode new information:

7

The bank robbers left in a rush. It was the youngest one who drove the van. (Wagner, 2012)

So the prominence in the CC is driven by context—or the preference about what context we accommodate when there is none. What about the RRC parse?

Restrictive relative clause Parse (RRC) 8

b. Who called?
It was ∼[the HUMORIST ___ **was leaving the SCENE**]$_F$

- The RRC-parse has a preference for prominence within the RC (usually at the end of)—**why?**
- Could RRCs simply tend to have a focus structure that ends up making the RC accented—so only the RC, not the head, is F-marked?

9

It was ∼[the humorist [that ___ was leaving the SCENE]$_F$]

- RRCs get accented when they a focus that answers an implicit question (here: *Which humorist?*)
- This prosody would imply that there were multiple (potential) humorists

A focus account of RRC prosody could she light why some RRCs have a different prosody:(Bresnan, 1972) argues that this depends on whether it's object- or subject-extracted, but this doesn't appear to matter in the example here:

10

A: What are you hiding in your hands?
B1: It's the FLY that ___ bothered you.
B2: It's the FLY that you complained about ___

- Prominence on *fly* seems preferable here—**unless there were multiple ones**
- So maybe prominence in the RRC case is also focus/context driven

This focus account could potentially also explain the RRC bias itself:

- The RRC-parse requires more unsupported assumptions about the context (multiple potential referents; a QUD to which we get no clues)
- The CC wears the QUD on its sleeve (or at least that's what we appear assume without a rich context, judging by the deaccentuation of the RC)

But... isn't this just another version of Altmann et al. (1992)'s account of an anti-RRC bias in terms of **unsupported presuppositions**?

And didn't we have reasons not to believe this? (Portner, 1989; Clifton Jr and Ferreira, 1989; Staub et al., 2018)

Here's the example that is often used against Altmann et al. (1992)'s presupposition story (Clifton Jr and Ferreira, 1989; Staub et al., 2018, i.a.). You can say the following without believing that there is more than one unicorn:

11

The unicorn that Richard saw was the only unicorn there ever was.

- But note this example can naturally have prominence on *unicorn*—unless we want to make salient other unicorns, fictional or imagined
- This is quite compatible with the focus account of RRC prosody (and a focus account of the bias)
- Were Altmann et al. (1992) maybe on the right track, iff slightly off target?

  The alternative explanation in Staub et al. (2018), cited in Guo et al., is not so plausible here: **Avoid unforced filler-gap dependencies**—*both* CC- and RRC-parse have a filler-gap dependency! Unless elided filler-gap dependencies also count...

What do the disambiguation effects here tell us about parsing and prominence?

A higher level conclusion:

- Suppose that prosodic prominence in these stimuli helps avoid garden paths even when no context is provided
- This would imply that when parsing, we don't just select parses based on how likely a given structure is...
- ...but also on how likely the focus alternatives are that the structure implies given the prominence pattern

I think this is quite intriguing, and goes beyond prior focus effects in sentence parsing (e.g. Braun and Tagliapietra (e.g. 2010); Gotzner et al. (e.g. 2016); Husband and Ferreira (e.g. 2016) on parsing focus particles or on focus and attachment Carlson et al. (2009); Carlson and Tyler (2017))

1. Nele Ots: Cross-linguistic survey of intonation planning: a cognitive approach

2. Guo et al.: The Garden Path Leading to Intonational Phonology

3. Ferreira: Prosody, Syntax, and Conversational Language

- As the previous two sections illustrated, it's impossible to talk about any aspect of sentence prosody without running into Fernanda's work, or that of her students (and in one case above her brother!)
- What makes her (and their) work so important is that it freely combines insights from different fields and methods
- And how else would we ever figure out how sentence prosody works? (e.g. by dissociating performance and competence effects as in Ferreira 1991)

A perennial question about sentence prosody:

- If prosodic constituent structure is sensitive to syntactic constituent structure, why not just have one?And if they are distinct, how do they relate, and why in this way?

12

a. [The woman [with [the winning bet]]] agreed to give us part of the money
b. [The woman [who [had won [the bet]]]] agreed to give us part of the money

- Do the number of right boundaries really matter here?
- There are some other possible explanations for the duration difference observed in Cooper and Paccia-Cooper (1980)

a. [The woman [with [the winning bet]]] agreed to give us part of the money
b. [The woman [who [had won [the bet]]]] agreed to give us part of the money

- We learned from Guo et al. that RRC parses are avoided
- Structure (b) is ambiguous—the RC could be restrictive or non-restrictive (as in: *Mary, who had won the bet, agreed to give us part of the money*)
- Could it be that some speakers took the RC to be non-restrictive? This would explain the durational difference (see duration effects in non-restrictive relative clauses see Poschmann and Wagner, 2016, i.a.)

Another possibility is that the increase in duration is a processing effect: the processing of the filler-gap dependency of the RRC (cf. Staub et al., 2018)

- Why be skeptical about counting the number of right parentheses? Because they often do not seem to matter
- Prosodic phrasing in English is constrained by syntactic constituency, but also very flexible
- The flexibility is sometimes arguably simply due to variation in the size of the planning window (as evidenced by pitch scaling in Ots's paper)

PP Attachment

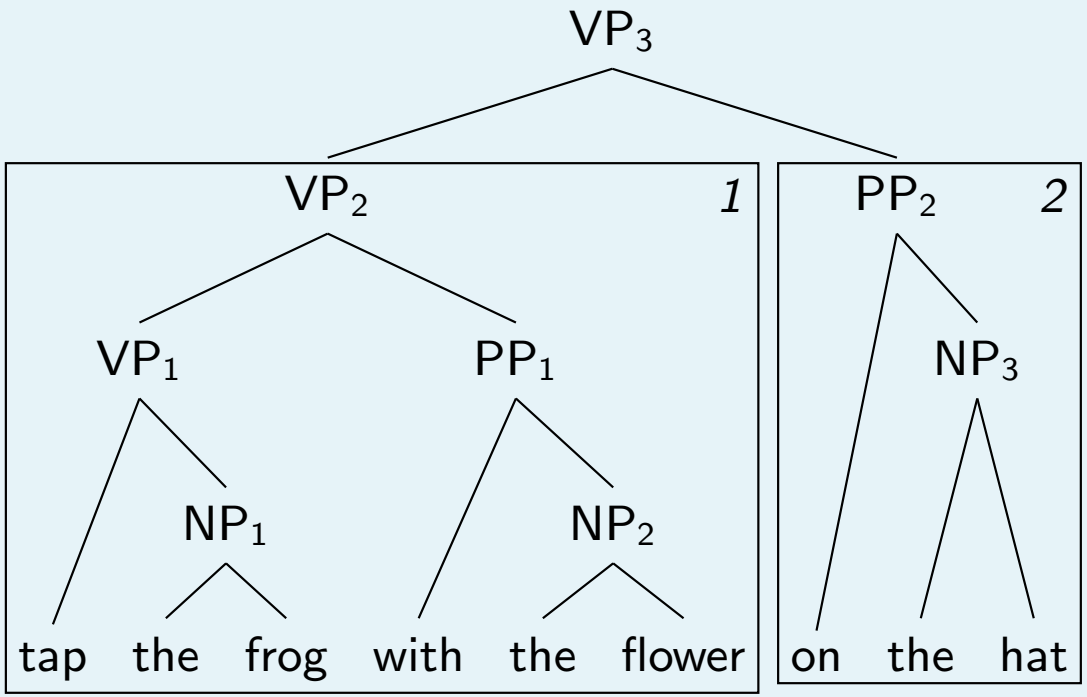
Hirsch and Wagner (2015, 2016): Variable phrasing in PP-attachment:

(1) Tap the frog with the flower on the hat.

- This sentence is has various readings, let's focus here on the one that can be paraphrased with *tap the frog, using the flower, and tap it on the hat*
- One way to pronounce this structure is without clear intonational boundaries.
- But there are other possibilities
- We can think of them as choices of how much structure to 'chew off' the tree at a time

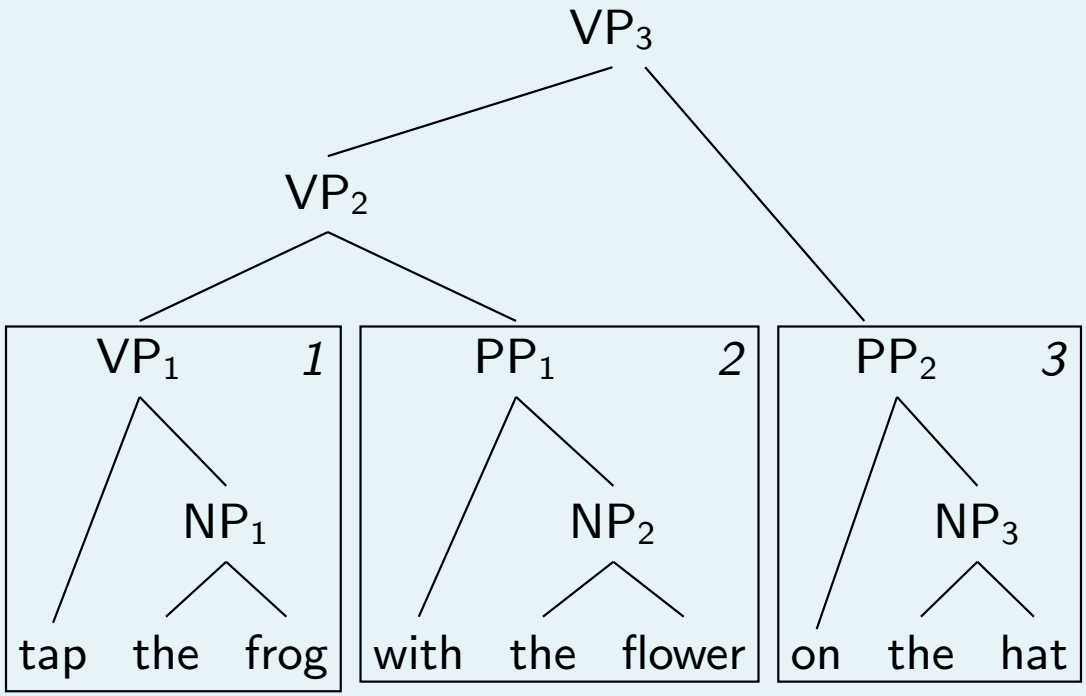
Example

Tap the frog with the flower ◇ on the hat.



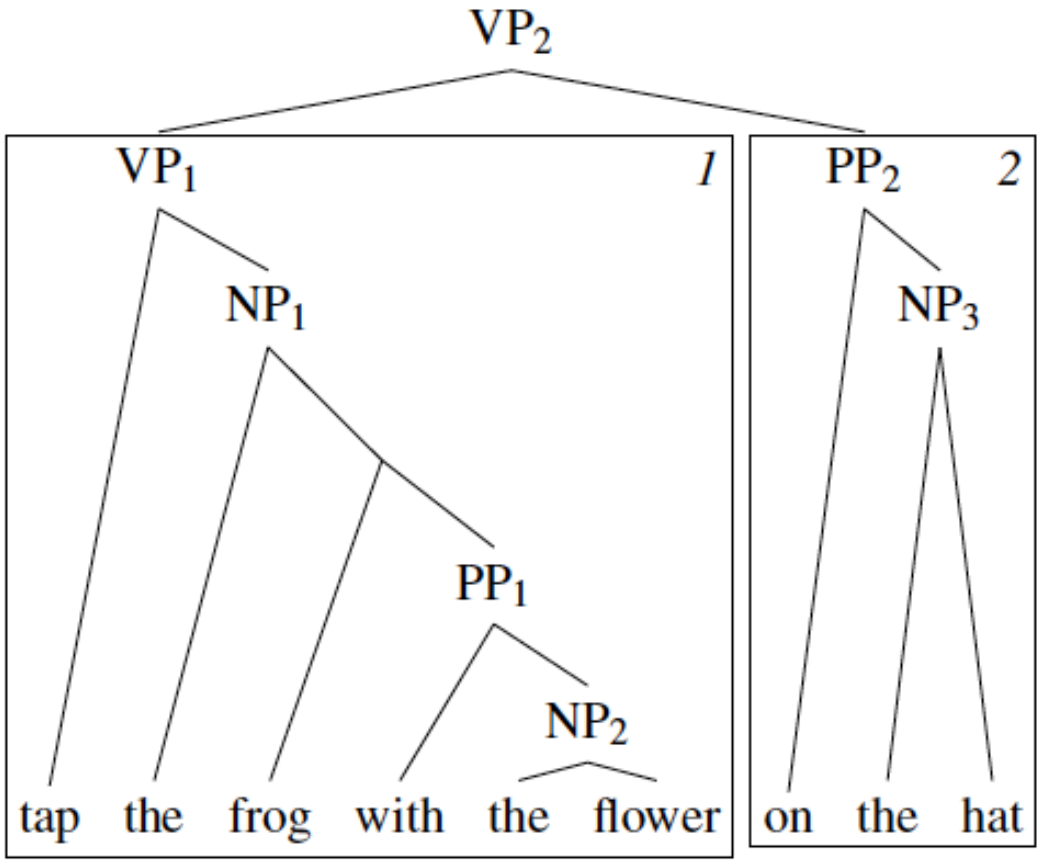
Example

Tap the frog ◇ with the flower ◇ on the hat.

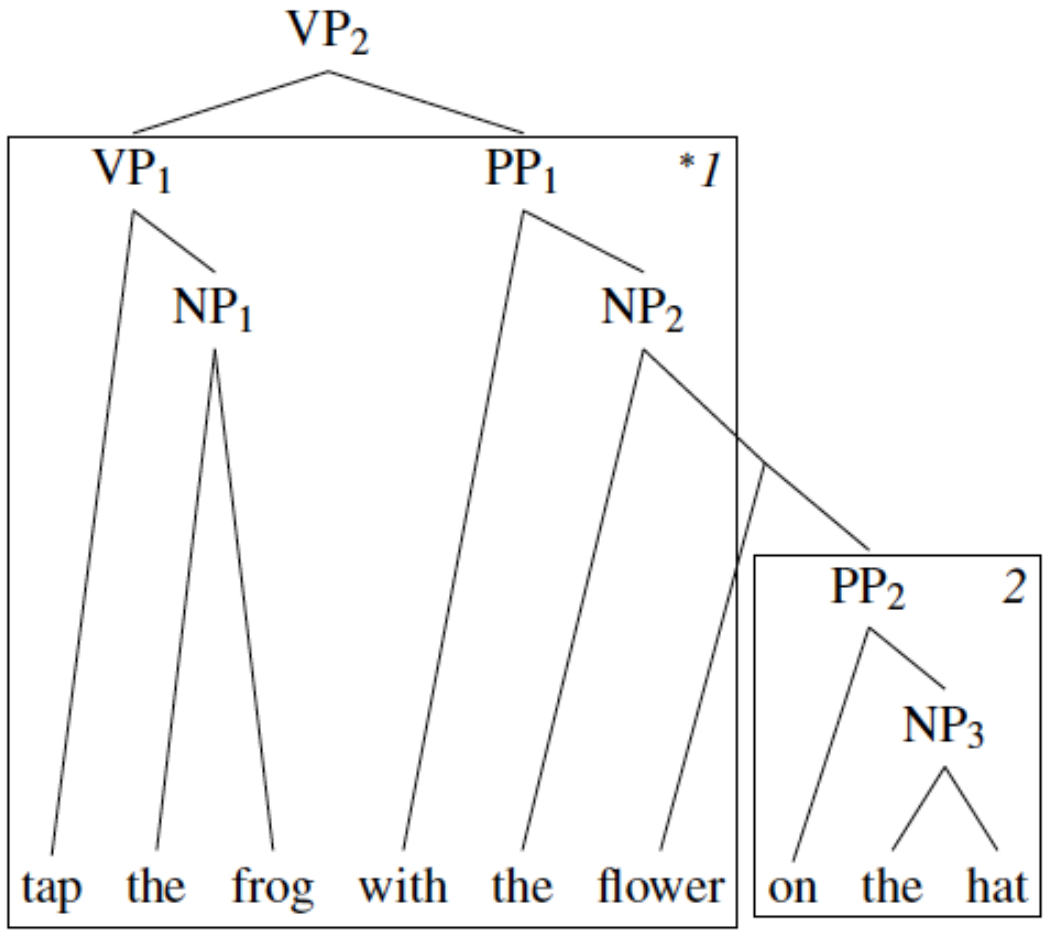


- How much a speaker will 'chew off' in a given planning chunk depends on many factors
- The predictability of upcoming material (Turk, 2008), individual differences in working memory (Martin et al., 2014; Swets et al., 2014)
- The complexity of the constituents (Wheeldon and Lahiri, 1997; Griffin, 2003; Meyer et al., 2007; Wheeldon, 2012) and Ots (2024)
- ...and task-related cognitive demands (Wagner et al., 2010; Swets et al., 2013) and Ots (2024)

- The most common choice for prosodic phrasing in our production data for this structure is after the first PP
- We argue that this much material is a sweet spot (and better than just planning *tap the frog*
- But this is actually **also** the most frequent realization of a more nested structure in which the first PP attaches to the first noun phrase, **despite a different number of right brackets**:



...while this phrasing is not attested for the right-branching structure, where planning that much material would violate syntactic constituency:



- This data makes sense of planning scope is variable, but constrained by syntactic constituency
- Of course, there do seem mismatches between syntax and prosody—here a famous example (thanks to Stefanie Shattuck-Hufnagel and Alice Turk for a recent discussion of this):

13

Sesame street is presented by ◇ the children's television network.
(Shattuck-Hufnagel and Turk, 1996)

- This constituency could be derived by right node raising—which here facilitates incremental structure building for dramatic effect)
- A recent case for revisiting syntactic assumptions behind mismatches is made in Royer (2020)'s work on phrasing in Mayan

But what about conversational language?

- As Fernanda observes, the messy, fragmented, and disfluent stuff we actually say conflicts with ideas about prosody based on constructed examples
- A promising perspective on this comes from another student of Fernanda's:

14

She liked the, um, the tomato.(Husband, 2015)

- Husband (2015) shows compelling evidence that repairs have the same syntactic properties as right-node-raising constructions

- If prosody contradicts syntax, maybe it's telling us something about syntax that syntacticians haven't been able to hear
- ...or wanted to hear, because it leads to a very different kind of syntax
- ...and might end up necessitating labphon studies to test hypotheses
- I personally think that what prosody is telling us is that **performance units/prosodic constituents** are **syntactic constituents**—often despite appearances

The prosody-syntax interface

Performance units = Syntax * Planning Constraints

- Admittedly, the syntactic theory that can truly deliver on this equation remains a promisory note—but to me it seems like something worth trying
- And of course it will be only as good as the syntactic and phonological predictions that it makes

Altmann, G. T., Garnham, A., and Dennis, Y. (1992). Avoiding the garden path: Eye movements in context. *Journal of Memory and Language*, 31(5):685–712.

Bock, J. and Levelt, W. (1994). Language production: Grammatical encoding. In Gernsbacher, editor, *Handbook of psycholinguistics*, pages 945–984. Academic Press, San Diego.

Braun, B. and Tagliapietra, L. (2010). The role of contrastive intonation contours in the retrieval of contextual alternatives. *Language and Cognitive Processes*, 25(7):1024–1043.

Bresnan, J. W. (1972). Stress and syntax: A reply. *Language*, 48(2):326–342.

Brown-Schmidt, S. and Konopka, A. E. (2014). Processes of incremental message planning during conversation. *Psychonomic bulletin & review*, 22(3):833–843.

Büring, D. (2016). *Intonation and meaning*. Oxford University Press.

Carlson, K., Dickey, M. W., Frazier, L., and Clifton Jr, C. (2009). Information structure expectations in sentence comprehension. *Quarterly Journal of Experimental Psychology*, 62(1):114–139.

Carlson, K. and Tyler, J. C. (2017). Accents, not just prosodic boundaries, influence syntactic attachment. *Language and Speech*, page 0023830917712282.

Clifton Jr, C. and Ferreira, F. (1989). Ambiguity in context. *Language and cognitive processes*, 4(3-4):SI77–SI103.

Cooper, W. E. and Paccia-Cooper, J. (1980). *Syntax and Speech*. Harvard University Press, Cambridge, Mass.

Ferreira, F. (1991). Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of Memory and Language*, 30(2):210–233.

Ferreira, F. and Swets, B. (2002). How incremental is language production? evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language*, 46:57–84.

Ferreira, V. S. (1996). Is it better to give than to donate? syntactic flexibility in language production. *Journal of memory and language*, 35(5):724–755.

Fuchs, S., Petrone, C., Krivokapić, J., and Hoole, P. (2013). Acoustic and respiratory evidence for utterance planning in german. *Journal of Phonetics*, 41(1):29–47.

Gotzner, N., Wartenburger, I., and Spalek, K. (2016). The impact of focus particles on the recognition and rejection of contrastive alternatives. *Language and Cognition*, 8(1):59–95.

Griffin, Z. (2003). A reversed word length effect in coordinating the preparation and articulation of words in speaking. *Psychonomic Bulletin and Review*, 10:603–609.

Hedberg, N. (2013). Multiple focus and cleft sentences. *Cleft structures*, pages 227–250.

Hirsch, A. and Wagner, M. (2015). Rightward movement affects prosodic phrasing. In Ozyildiz, D. and Bui, T., editors, *Proceedings of the 45th Meeting of the North-East Linguistic Society (NELS)*.

Hirsch, A. and Wagner, M. (2016). Variable prosodic phrasing in a theory of incremental production planning. Ms.

Husband, E. M. (2015). Self-repairs as right node raising constructions. *Lingua*, 160:20–37.

Husband, E. M. and Ferreira, F. (2016). The role of selection in the comprehension of focus alternatives. *Language, Cognition and Neuroscience*, 31(2):217–235.

Konopka, A. E. (2012). Planning ahead: How recent experience with structures and words changes the scope of linguistic planning. *Journal of Memory and Language*, 66(1):143–162.

Lindsley, J. R. (1975). Producing simple utterances: How far ahead do we plan? *Cognitive Psychology*, 7(1):1–19.

Martin, R. C., Yan, H., and Schnur, T. T. (2014). Working memory and planning during sentence production. *Acta psychologica*, 152:120–132.

Meyer, A. S., Belke, E., Häcker, C., and Mortensen, L. (2007). Use of word length information in utterance planning. *Journal of Memory and Language*, 57(2):210–231.

Myachykov, A., Scheepers, C., Garrod, S., Thompson, D., and Fedorova, O. (2013). Syntactic flexibility and competition in sentence production: The case of english and russian. *Quarterly Journal of Experimental Psychology*, 66(8):1601–1619.

Portner, P. (1989). Processing indefinite noun phrases in quantified sentences. University of Massachusetts, Amherst.

Poschmann, C. and Wagner, M. (2016). Relative clause extraposition and prosody in German. *Natural Language & Linguistic Theory*, 34(3):1021–1066.

Royer, J. (2020). An apparent syntax-prosody mismatch as right extraposition: Evidence from mayan. In Whitmal, A., Song, Y., and Asatryan, M., editors, *NELS 50: Proceedings of the Fiftieth Annual Meeting of the North East Linguistic Society*, volume 3, pages 113–126.

Shattuck-Hufnagel, S. and Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, 25(2):193–247.

Staub, A., Foppolo, F., Donati, C., and Cecchetto, C. (2018). Relative clause avoidance: Evidence for a structural parsing principle. *Journal of Memory and Language*, 98:26–44.

Swets, B., Desmet, T., Hambrick, D. Z., and Ferreira, F. (2007). The role of working memory in syntactic ambiguity resolution: A psychometric approach. *Journal of Experimental Psychology: General*, 136(1):64.

Swets, B., Jacovina, M. E., and Gerrig, R. J. (2013). Effects of conversational pressures on speech planning. *Discourse Processes*, 50(1):23–51.

Swets, B., Jacovina, M. E., and Gerrig, R. J. (2014). Individual differences in the scope of speech planning: evidence from eye-movements. *Language and Cognition*, 6(01):12–44.

Turk, A. (2008). Prosodic constituency signals relative predictability. Ms. University of Edinburgh.

Wagner, M. (2012). Focus and givenness: A unified approach. In Neeleman, A. and Kučerová, I., editors, *Contrasts and Positions in Information Structure*, pages 102–147. Cambridge University Press, Cambridge.

Wagner, V., Jescheniak, J. D., and Schriefers, H. (2010). On the flexibility of grammatical advance planning during sentence production: effects of cognitive load on multiple lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(2):423.

Wheeldon, L. (2012). Producing spoken sentences: The scope of incremental planning. In Fuchs, S., Weirich, M., Pape, D., and Perrier, P., editors, *Speech production and perception: Vol. 1. Speech planning and dynamics*, pages 97–118. Peter Lang, Frankfurt, Germany.

Wheeldon, L. and Lahiri, A. (1997). Prosodic units in speech production. *Journal of Memory and Language*, 37:356–381.

Williams, E. (1997). Blocking and anaphora. *Linguistic Inquiry*, 28(4):577–628.