Speech planning: key models and results, part I

1. Influential Levelt & colleagues model

- Top references
  - Levelt 1993: book-length presentation
  - Levelt, Roelofs & Meyer 1999: updated presentation, plus computational implementation, **weaver++** (figures below from here)

1.1 Lexical representations

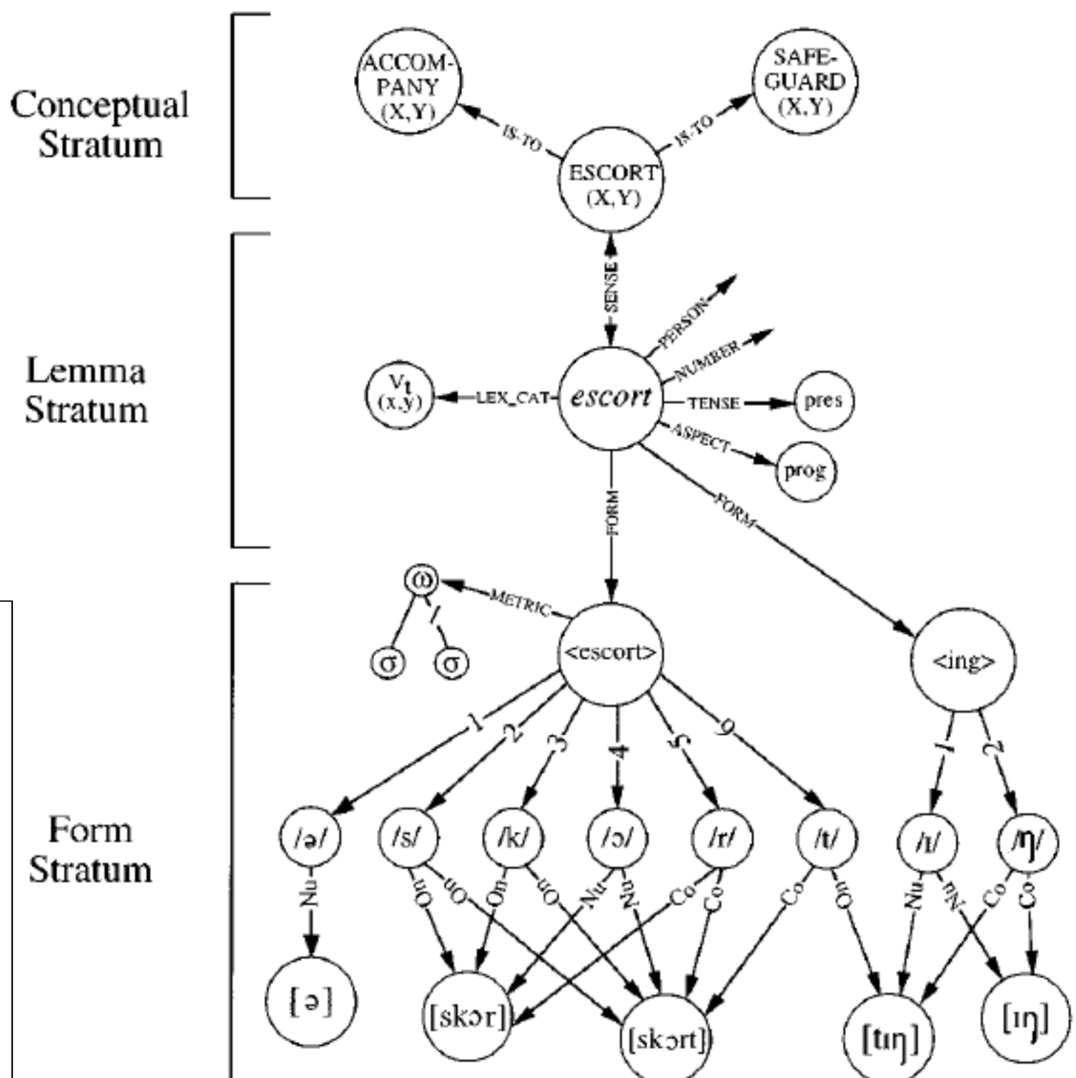
- Are complex, exist at many levels

no phonology here!

still no phonology!

these entries actually contain phonological info (which prosody, and segments to access)

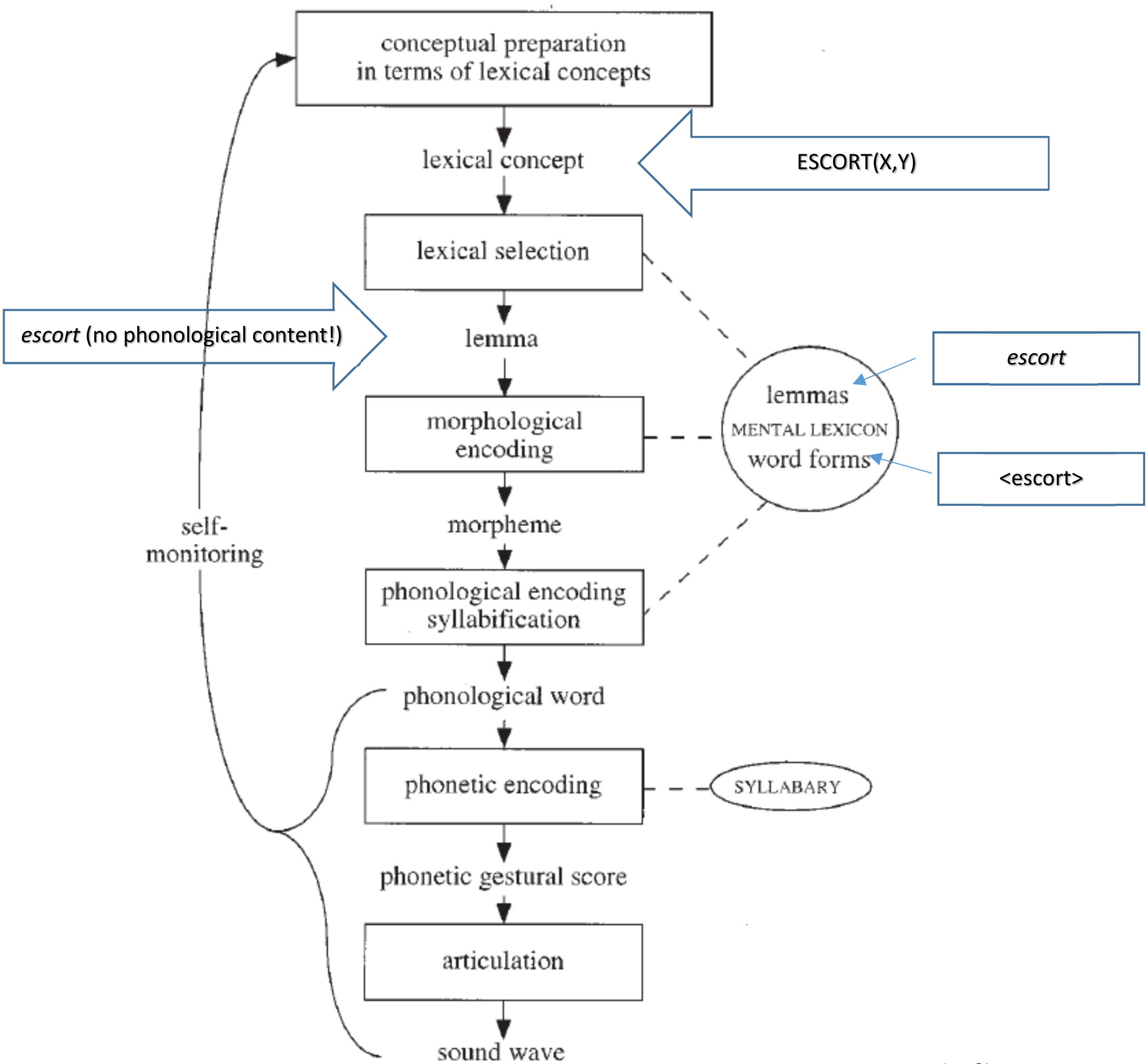
notice that prosodic shape is separate from segmental content. Which segments belong to which syllable is not stored, but computed online.



(p. 4)

1.2 Producing a word

- Also requires many stages (arrows mine)



(p. 3)

2. Classic findings about these stages

2.1 Speech errors that swap words can be longer-distance than speech errors that swap sounds

- *The test will be about discussing the class* vs. *thollow hud* (Dell & O'Seaghdha 1992)
  - interpretation: lemma for *class* is ready well in advance
    - and gets put into the wrong slot
  - word-form entry, or at least segmental content, for <thud> is available only a little bit in advance
    - and likewise, part of it gets put into the wrong slot

2.2 Importance of clause boundaries (overview from Dell & O'Seaghdha 1992)

- Word-swap errors are usually in the same clause (Garrett 1975)

- Hesitations tend to be at clause boundaries (Holmes 1988)

2.3 Smith & Wheeldon 2004: semantically related nouns can interfere at longer distances than phonologically related nouns

- Method
  - See picture and word, with motion: [1]  AXE
  - Say *The saw and the axe move down, The saw moves towards the cat,* etc.
  - Dependent variables: how long to start talking, error rate

- Result
  - If picture and word are semantically related (saw/axe), takes longer to start talking
    - true in both structures: [*the* ***saw*** *and the* ***axe***] *move down* or [*the* ***saw***] *moves towards* [*the* ***axe***]
    - but weaker effect if not in same phrase
    - interpretation: both words' lemmas can be accessed before speech starts, especially if they're in the same noun phrase
  - If picture and word are phonologically related and in same noun phrase (*the f**lag** and the b**ag***), faster and fewer errors
    - but only if same at end (*flag/bag*), not beginning (*cat/cap*)—they're not sure why
  - No effect if phonologically related and far apart (*the **cat** moves above the **cap***)
    - interpretation: before the utterance starts, phonological information from both nouns in the subject NP can be available, but not from a noun later in the sentence

[1] Thanks, thenounproject.com

2.4 Dell & O'Seaghdha 1992: "lemmas are buffered […] before they are phonologically specified"

- Method
  - See vaguely logic-looking formula: *REMOVE(BOXER, COAT)* or *REMOVE(BY BOXER, COAT)*
  - Prepare utterance: *The boxer removed the coat* or *The coat was removed by the boxer*
  - Filler trials: see *, say utterance
  - Critical trials: see target word (*COAT* or *COAL* or *SHIRT*) and read it aloud
  - Dependent variable: response time in saying target word

- Result
  - *The* ***coat*** *was removed by the boxer* slows down response to *COAL*
    - interpretation: *coat* is already phonologically accessed, enough to compete with planning *coal*
  - *The boxer removed the* ***coat*** actually speeds up response to *COAL*!
    - interpretation: maybe just the /k/ or so has been accessed, and/or the prosodic shape, which are helpful for *coal*
  - Effect of semantically related prime, *SHIRT*, is more complicated (there's another experiment)
    - but basically, it doesn't much depend on whether *coat* was early or late in the prepared sentence
- Interpretation: again, something like lemma access happens earlier than full phonological access

3. Size of the planning unit: looking up lemma or so

3.1 ("phrasal-level") prosodic word (Wheeldon & Lahiri 1997; Wheeldon & Lahiri 2002)

Dutch

- 1997: Syntactic words vs. prosodic words vs. syllables
  - p-words or lexical words matter, not syntactic words or syllables

	(ik zoek het) (water)	*(ik zoek) (water)*	*(ik zoek) (vers) (water)*
	'I seek the water'	'I seek water'	'I seek fresh water'
	2 p-words	**2 p-words**	**3 p-words**
	2 lexical words	**2 lexical words**	**3 lexical words**
	4 syntactic words	3 syntactic words	4 syntactic words
	5 syllables	4 syllables	5 syllables
"naming latency": how fast it takes the participant to start talking	faster (just about identical) although *ik zoek water* has shorter <u>duration</u>		slower

- follow-up experiment: no, it's p-words that matter, not lexical words

	(ik zoek)	*(ik zoek het) (water)*	*(ik zoek) (het)*	*(ik zoek) (vers) (water)*
	'I seek'	'I seek the water'	'I seek it'	'I seek fresh water'
	1 p-word	**2 p-words**	**2 p-words**	**3 p-words**
	1 lexical word	2 lexical words	1 lexical word	3 lexical words
	2 syntactic words	4 syntactic words	3 syntactic words	4 syntactic words
	2 syllables	5 syllables	4 syllables	5 syllables
naming latency	fastest	medium (just about identical)		slowest

- 2002: Oh no, here it looks like syntactic word matters!

	óoglìd	*órgel*	*orkáan*	*òud líd*
	'eyelid'	'organ'	'hurricane'	'old member'
	2 p-words	1 p-word	1 p-word	2 p-words
	1 syntactic word	**1 syntactic word**	**1 syntactic word**	**2 syntactic words**
naming latency	faster (all about the same)			slower

  - their interpretation: actually, it's the "phonological word at the phrasal level", some kind of "superword"
    - I buy this:
      - by some diagnostics, Dutch compounds are 2 p-words (syllabification)
      - but still they have a single primary word stress, unlike a phrase, which has two word stresses

3.2 Maybe it's flexible (Wagner, Jescheniak & Schriefers 2010—different Wagner!)

- Note: this is about **access, not encoding** (we'll return to this point and all the terminological confusion next time)

- Different experimental setup, different results
  - simple sentence: two nouns
    - *the frog is next to the mug*
  - more-complex sentence: mostly just the first noun
    - *the blue frog is next to the blue mug*
  - interference from additional task, or utterance variety: first noun only
    - participant has to switch between simple and complex format, depending on whether the objects they see depicted are (in nature) small or big
    - *the (blue) frog is next to the (blue) mug*
  - interference from working-memory task: back to both nouns
    - before each trial, participants given 5 digits or 5 adjectives to memorize
    - *the frog is next to the mug*
    - Their speculation: this task is "not directily related to utterance production" (p. 435)

- Experimental setup
  - Subject sees side-by-side drawings[2]  
  - Has to describe them, in format *the __ is next to the __*
  - Headphones play a distractor word
    - if *toad*, should interfere with first word
    - if *cup*, should interfere with second word
  - Dependent variables
    - how long it takes to start talking (compared to semantically unrelated distractor word)
    - error rate
  - If only *toad* harms performance, only first word was accessed before participant started talking
  - If both *toad* and *cup* harm performance, both words were already accessed

- Also, differences between fast responders and slow responders
    - Slow responsers plan more before starting to talk (in simple task)

I think that's enough of this for one day—let's move on to…

4. Comments on how the two handbook articles relate to previous articles
(Goldrick 2014; Buchwald 2014)

4.1 Wagner 2012 (English *–ing/-in'*, research program; Kie)

- In Wagner's *–ing/-in'* experiment, the following word could be *the* (starts with coronal: encourages *–in'*) or *a* (doesn't start with coronal)
  - But there's a coarser difference too
    - *the* = CV (could start with coronal)
    - *a* =V (can't possibly be coronal)
  - **Goldrick** reviews evidence that a word's CV structure is retrieved separately from its segmental structure.
    - plausibly, CV structure is retrieved earlier than segmental
    - fact that *a* = V (and therefore doesn't start with a coronal) could be available earlier than the fact that *the* starts with a coronal
  - This predicts a bigger effect for *the* vs. *a* than for *the* vs. *my*, especially in the no-clause-boundary condition
    - Would be interesting to replicate experiment with third condition (*my*)
    - across clause boundary, we might expect difference between *the* and *my* to become especially small (I will draw hypothetical plot on board)

[2] Thanks, emojidex

- **Buchwald** emphasizes the difference between phonological encoding, which outputs something like /kæt/, and the phonetic processing, which fills in things like aspiration (i.e., postlexical)
  - This makes me wonder about the processes Wagner discusses, which all cross phoneme boundaries
    - *–ing/-in'*, tone sandhis, French liaison
  - Maybe the phonetic processor is able to do these things
    - but especially in the case of *–ing/in'*, which requires selecting between allomorphs, maybe phonological encoding needs to already do the work
  - This might mean that these types of morphology/phonology might pattern differently w.r.t. planning than truly postlexical stuff like English tapping

4.2 Kilbourn-Ceron, Wagner & Clayards 2016 (English tapping; Meng)

4.3 Kilbourn-Ceron & Sonderegger 2018 (Japanese high V devoicing; Canaan)

4.4 Kilbourn-Ceron 2017b (French liaison; Jesse if he wants to)

4.5 Tanner, Sonderegger & Wagner 2017(English t/d deletion; Isabelle)

4.6 Tamminga 2015 (English t/d deletion; Brice)

4.7 Gahl & Garnsey 2004 (English t/d deletion; Allie)

4.8 MacKenzie 2012, ch. 5 (English is/’s, has/’s, will/’ll; Beth if she wants to)

4.9 MacKenzie 2016 (English is/'s; Jacob)

4.10 Lamontagne & Torreira 2017 (Spanish V hiatus; Kie)

- **Goldrick**'s opening figure shows a word's syntactic properties being accessed early
  - That should include gender
  - In a Spanish phrase like *buena estrella* 'good star, lucky star'…
    - gender of *estrella* must be known before final vowel of *buena* can be planned
  - Whereas in *suegra ejemplar* 'exemplary mother-in-law'…
    - nothing syntactic about *ejemplar* needs to be known before final V of *suegra* can be planned
  - In fluent productions of *buena estrella*, *estrella*'s syntactic planning needs to be done in good time, so <u>maybe</u> its phonological planning is also a little more advanced?
  - Probably not big enough to see an effect
    - but I guess there's a hypothesis to be tested about pairs where Word1 agrees with Word2 (and Word2 is the head, driving the agreement) vs. other bigrams
  - By the way, Lamontagne, in other slides, does find that feminine suffix –*a* (*herman-a* 'sister') is deleted more than other final –*a* (*ahora* 'now')
    - but doesn't distinguish whether the –*a* is in the head, is before the head it's agreeing with, or is after the head it's agreeing with

- **Buchwald** discusses "the role of the syllable in phonetic processing"
  - Relevant to the Spanish case, because vowel deletion changes the syllabification
    - [es.**te.a**.moɾ] vs. [es.**ta**.moɾ]
    - [nues.**tra.es**.kue.la] vs. [nues.**tras**.kuela]
  - Buchwald reviews evidence that high-frequency syllables have processing advantages
    - but says it's unclear whether these advantages happen during phonological encoding or phonetic processing (or even later)
    - If the frequency of syllables like [ta] and [tɾas] matters in predicting Spanish deletion…
      - then I think that tells us that this really is the syllabification, and we're not just looking at some kind of very late blending/competition of the two vowel gestures

5. Up next:

- I'll present some more of these highlights from the general speech-planning literature

- **Keating & Shattuck-Hufnagel 2002**
- **Wheeldon, Meyer & Smith 2006** plus **Wheeldon 2013**
  - Here's a suggestion for a different task for these two: let's each bring one highlight to share about the paper:
    - something you didn't know before reading it
    - what stood out to you the most
    - your biggest question about the paper
    - …?

- Kie: present highlights from speech error literature (as relevant to OCP!)
  - I think all this will take through the end of next week—we can divide up the next set of readings of Tuesday

- On the horizon: I have a group writing exercise in mind (maybe 1 hour, maybe 2)

References

Buchwald, Adam. 2014. Phonetic processing. In Matthew Goldrick, Victor Ferreira & Michele Miozzo (eds.), *The Oxford handbook of language production*, 245–258. Oxford: Oxford University Press.

Dell, Gary S. & Padraig G. O'Seaghdha. 1992. Stages of lexical access in language production. *Cognition* 42(1–3). 287–314. doi:10.1016/0010-0277(92)90046-K.

Gahl, Susanne & Susan Marie Garnsey. 2004. Knowledge of Grammar, Knowledge of Usage: Syntactic Probabilities Affect Pronunciation Variation. *Language* 80(4). 748–775. doi:10.1353/lan.2004.0185.

Garrett, M. F. 1975. The Analysis of Sentence Production. In Gordon H. Bower (ed.), *Psychology of Learning and Motivation*, vol. 9, 133–177. Academic Press. doi:10.1016/S0079-7421(08)60270-4. http://www.sciencedirect.com/science/article/pii/S0079742108602704.

Goldrick, Matthew. 2014. Phonological processing: The retrieval and encoding of word form information in speech production. In Matthew Goldrick, Victor Ferreira & Michele Miozzo (eds.), *The Oxford handbook of language production*, 228–244. Oxford: Oxford University Press.

Holmes, V. M. 1988. Hesitations and sentence planning. *Language and Cognitive Processes* 3(4). 323–361. doi:10.1080/01690968808402093.

Keating, Patricia & Stefanie Shattuck-Hufnagel. 2002. A prosodic view of word form encoding for speech production. *UCLA Working Papers in Phonetics* 101. 112–156.

Kilbourn-Ceron, Oriana. 2017. Speech production planning affects phonological variability: a case study in French liaison. In Karen Jesney, Charlie O'Hara, Caitlin Smith & Rachel Walker (eds.), *Proceedings of the 2016 Annual Meeting on Phonology*.

Kilbourn-Ceron, Oriana & Morgan Sonderegger. 2018. Boundary phenomena and variability in Japanese high vowel devoicing. *Natural Language and Linguistic Theory*. doi:10.1007/s11049-017-9368-x.

Kilbourn-Ceron, Oriana, Michael Wagner & Meghan Clayards. 2016. The effect of production planning locality on external sandhi: a study in /t/. *Proceedings of the 52nd Annual Meeting of the Chicago Linguistic Society*.

Lamontagne, Jeffrey & Franciso Torreira. 2017. Production planning and directionality in external sandhi (poster). *Annual Meeting on Phonology*. New York University.

Levelt, Willem J. M. 1993. *Speaking: from intention to articulation*. MIT Press.

Levelt, Willem J. M., Ardi Roelofs & Antje S. Meyer. 1999. A Theory of Lexical Access in Speech Production. *Behavioral and Brain Sciences* 22(01). 1–38.

MacKenzie, Laurel. 2012. Locating variation above the phonology. University of Pennsylvania PhD dissertation.

MacKenzie, Laurel. 2016. Production planning effects on variable contraction in English. *Penn Working Papers in Linguistics* 22(2). 121–130.

Smith, Mark & Linda Wheeldon. 2004. Horizontal Information Flow in Spoken Sentence Production. *Journal of Experimental Psychology: Learning Memory and Cognition*. 675–686.

Tamminga, Meredith. 2015. Modulation of the following segment effect on coronal stop deletion [slides]. *NWAV 44*. Toronto.

Tanner, James, Morgan Sonderegger & Michael Wagner. 2017. Production planning and coronal stop deletion in spontaneous speech. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 8(1). doi:10.5334/labphon.96. http://journal-labphon.org/articles/10.5334/labphon.96/.

Wagner, Michael. 2012. Locality in phonology and production planning. In A McKillen & J Loughran (eds.), *Proceedings of the Montreal-Ottawa-Toronto (MOT) Phonology Workshop 2011. Phonology in the 21st Century: In honour of Glyne Piggott. McGill Working Papers in Linguistics 22(1).*

Wagner, Valentin, Jörg D. Jescheniak & Herbert Schriefers. 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). 423–440. doi:10.1037/a0018619.

Wheeldon, Linda. 2013. Producing spoken sentences: the scope of incremental planning. In S Fuchs, M Weirich, D Pape & P Perrier (eds.), *Speech production and perception volume 1: speech planning and dynamics*, 93–114. Peter Lang.

Wheeldon, Linda & Aditi Lahiri. 1997. Prosodic Units in Speech Production. *Journal of Memory and Language* 37(3). 356–381. doi:10.1006/jmla.1997.2517.

Wheeldon, Linda R. & Aditi Lahiri. 2002. The minimal unit of phonological encoding: prosodic or lexical word. *Cognition* 85(2). B31–B41. doi:10.1016/S0010-0277(02)00103-8.

Wheeldon, Linda R, Antje S Meyer & Mark Smith. 2006. Language Production, Incremental. *Encyclopedia of Cognitive Science*. John Wiley & Sons, Ltd. doi:10.1002/0470018860.s00252. http://onlinelibrary.wiley.com/doi/10.1002/0470018860.s00252/abstract.