Summary: Language production and dialogue

So far, we have talked a good deal about the mental processes required to comprehend language – but we also often engage in the act of *producing* language. Indeed, in our very first week, we talked about how one of the hallmarks of language was how productive it is – even given the constraints of language (e.g., those posed by syntax), we can still produce entirely novel utterances.

Language production can be thought of as the inverse of language comprehension – rather than mapping from sounds to meaning, we first access the meaning we want to convey (**conceptualization**), plan which words and sounds we need to say (**formulation**), and then actually execute our plan (**articulation**).

That these steps occur in this order may seem intuitive; it's also supported experimentally. For instance, if a speaker has to name two pictures (e.g., *church* followed by *candy*), hearing a semantic associate for either picture (*mosque* or *sweet*) slows down production (because the speaker has to resolve competition between related word candidates). Hearing a phonological associate for the first picture (*churn*) facilitates production (because the correct phonemes are activated by both words), but hearing a phonological associate for the second picture (*candle*) does not affect production times. Taken together, these data suggest that meaning is conceptualized before sounds are planned, as later utterances are affected by semantic associates but not by phonological ones.

That these steps proceed in this set order is also evidenced by **speech errors**; exchange errors (where two linguistic units in an utterance are interchanged, such as *You have tasted the whole worm* instead of *You have wasted the whole term*) suggest that the upcoming sounds/words are formulated before current ones are articulated.

Indeed, much of our knowledge about the language production system comes from speech errors, since the structure of the system needs to also account for the sorts of errors we see (as well as the kinds we don't). For instance, speech errors suggest that lexical entries exist in base (uninflected) forms called **lemmas**, and that inflections (e.g., tense markers) are applied after selection of the **lemma**. This would explain why people make errors like *I'd hear it if I knew it* (instead of *I'd know it if I heard it*) – note that such an error involves an exchange of lemmas (*hear* and *know*) but correct application of tenses (base form for the first verb, past tense for the second). This also explains why we wouldn't see an error like *I'd heard it if I know it*.

Finally, substitution errors (where an intended unit is swapped out for an unintended one, as in *Please turn in your pencils* instead of *Please turn in your papers*) suggest that partially activated candidates (e.g., at the level of meaning) can influence activation at subsequent levels of production. This means that there is **cascading activation** from one level to the next; that is, the three steps of speech production are not modular (that is, they are not isolated from / independent of each other).

Further evidence for cascading activation comes from the fine-grained phonetics of our speech output. For instance, we see that speakers hyperarticulate words that are perceptually confusable with other words. The voice-onset time (VOT) of /k/ in a word like *cod* is relatively long, which makes it more distinct from the phonologically similar (and therefore partially activated) word

God. However, the VOT of /k/ in the word *cop* is relatively short, since there is no competitor (**gop*) that the speaker needs to contrast it with. This suggests that competitor words, even though not produced, influence the articulation of intended words. Interestingly, this hyperarticulation effect occurs even if the competitor word (e.g., God) is not contextually relevant, suggesting that this effect does not arise because of listener-oriented process (i.e., because the speaker is trying to make things easier *for the listener*) but because of some speaker-internal set of processes.

However, there are several aspects of production that the speaker modifies for the benefit of the talker (**audience design**). Evidence suggests that adults modify their productions based on what information is held in **common ground** (shared between interlocutors), and listeners adjust their processing of speech depending on who is talking. These sorts of adjustments require some knowledge that other people can have different mental states (**theory of mind**), which children usually begin to acquire around the age of 4.