Title
A Mechanistic Framework for Explaining Audience Design in Language Production.

Journal
Annual review of psychology, 70(1)

ISSN
0066-4308

Author
Ferreira, Victor S

Publication Date
2019

DOI
10.1146/annurev-psych-122216-011653

Peer reviewed

A Mechanistic Framework for Explaining Audience Design in Language Production

Victor S. Ferreira

Department of Psychology and Center for Research in Language

University of California, San Diego

Short title: A Mechanistic Audience Design Framework

Address for correspondence:

Victor S. Ferreira

Department of Psychology 0109

University of California, San Diego

9500 Gilman Drive

La Jolla, CA 92093-0109

USA

This research was supported by NIH HD051030.

Abstract

Audience design refers to when speakers fashion their utterances so as to cater to the needs of their addressees. Here, a range of audience-design effects are reviewed as organized by a novel cognitive framework for understanding audience-design effects. Within this framework, feedforward (or “one-shot”) production is responsible for *feedforward audience-design* effects, effects based on already known properties of the addressee (e.g., child vs. adult status) or a message (that it includes meanings that might be confusable). Then, a forward-modeling approach is described, whereby speakers independently generate communicatively relevant features to predict potential communicative effects. This can explain *recurrent processing audience-design* effects, effects based on features of the produced utterance itself, or idiosyncratic features of the addressee or communicative situation. Predictions from the framework are delineated.

Table of Contents

Keywords: Audience design, language production, forward modeling, prediction

Introduction

Most of the time, a speaker's primary reason for producing an utterance is to be understood by an addressee. A speaker may wish to transmit information ("It's raining."), to enlist cooperation ("Can you help me move this table?"), or to socialize ("How do you do?"). But, whatever the motivation for producing an utterance, it will only be fulfilled if addressees understand speakers' utterances well enough to know what the speaker intended to convey.

This raises the phenomenon in the field of the cognitive psychology of language that is termed *audience design*. That is, when choosing a set of linguistic features for an utterance, speakers must of course choose the words that convey their intended meanings ("it's raining" and "it's snowing" convey different intended meanings), and they must organize their words into sentence structures that convey who did what to whom ("dog bites man" and "man bites dog" convey different event relationships). But beyond such information-transmission functions, speakers produce some linguistic features to make their utterances easier to understand, and so more likely to be understood as well. When speakers do this, they engage in audience design.

For example, consider a sentence type known as a *sentence-complement structure*, such as, "The voter believed that the presidential candidate was lying." Whenever a speaker produces a sentence-complement structure, she or he must make a choice as to whether to include or omit an optional "that," as in (compared to the above), "The voter believed the presidential candidate was lying." These two sentences convey seemingly identical meanings, as the "that" itself conveys little or no information. In turn, this suggests that the choice to mention or omit the "that" is at least sometimes compelled by factors other than the meaning to be conveyed. One reasonable candidate factor is audience design.

Indeed, there is at least one good reason why the optional “that” in a sentence-complement structure might be used for audience-design purposes. Some sentences are what are called *garden-path* sentences. This refers to when a sentence invites one compelling interpretation (because it is especially simple or frequent) that is actually incompatible with the ultimate interpretation of the sentence. Without the “that,” sentence-complement structures can be garden-path sentences. Consider, “The voter believed the presidential candidate was lying.” Part-way through the sentence, after the noun phrase “the presidential candidate,” the sentence is compatible with a simple transitive analysis with “voter” as subject, “believed” as verb, and “presidential candidate” as direct object: “The voter believed the presidential candidate.” Ultimately, however, “presidential candidate” is not a direct object, but is instead the subject of the upcoming verb (an *embedded subject*), “was lying.” And so such a sentence “leads the comprehender down the garden path,” in that it initially invites one interpretation (“presidential candidate” as direct object as part of a simple transitive sentence), but ultimately requires a different analysis (“presidential candidate” instead as an embedded subject). At least in reading, such sentence-complement garden paths disrupt comprehension (see F. Ferreira & Clifton, 1986, Trueswell, Tanenhaus, & Garnsey, 1993).

Interestingly, with the “that,” sentence-complement structures are not garden-path sentences. That is, after the noun phrase “presidential candidate,” “The voter believed that the presidential candidate” is nearly impossible to interpret with “presidential candidate” as a direct object. This is because the “that” is part of a structure that includes the upcoming material as an embedded clause, in turn implying that the upcoming noun phrase is the embedded subject of that embedded clause. And so, it may be good if speakers avoid the potentially troublesome garden path, by mentioning the “that” and so eliminating the possibility of a direct-object interpretation. In turn, this would be an act of audience design: The speaker will have chosen the “that” to assist the comprehender to more efficiently understand his or her utterance.

A number of studies have investigated whether in fact, speakers avoid producing garden-path sentences. Although two studies have provided corpus-based evidence that writers may indeed craft their utterances to avoid garden paths (Elsness, 1984; Temperley, 2000), data from both spoken production experiments and conversational speech corpora suggest that speakers do not avoid garden paths (Arnold et al, 2004; Ferreira & Dell, 2000; Ferreira & Hudson, 2011; Ferreira & Schotter, 2013; Jaeger, 2010, 2011). Ferreira (2008) provides a framework for understanding why. Specifically, speakers may not avoid garden paths in part because doing so is likely very difficult, as it would require the production system to diagnose the garden path in the first place (a challenge that is discussed further below), and then determine an available remedy that itself doesn’t make the sentence even harder to comprehend (though the mention of the “that’ is a straightforward way to avoid garden paths in sentence-complement structures, there are many other types of garden paths, each potentially with different and often less obvious remedies). What’s more, whereas garden paths are demonstrably disruptive (as the long literature on garden-path comprehension shows), the degree of disruption they engender may not be worth speakers’ effort to try to avoid, given all of the other challenges that confront utterance production and comprehension in real time (see also Jaeger, 2006; Fukumura, 2018).

But the avoidance of garden paths is just one type of audience design that speakers might engage in. The goal of the remainder of this review is to organize audience design effects through a theoretical framework that the language production system may implement.

Feedforward Language Production

The present framework for understanding audience design begins with a general consensus account of how typical *feedforward* language production operates (e.g., Bock & Levelt, 1994; Dell, 1986; Ferreira & Slevc, 2007; Garrett, 1975; Levelt et al., 1999). "Feedforward" here does *not* mean the processing dynamic that contrasts with *feedback* processing (e.g., Dell, 1986 vs. Levelt et al., 1991); rather, "feedforward" instead refers to one-shot, first-pass production, in contrast with the type of recurrent processing that is elaborated below. The idea is that the act of production begins with a candidate utterance that is generated by standard, relatively uncontroversial language-production mechanisms.

Such a framework, elaborated somewhat to account for important aspects of audience design, is shown in Figure 1. Generally, an act of production begins with a communicative intention, that is, the goal that the speaker wishes to achieve by producing the utterance. For example, imagine a university professor who has interviewed candidates for her Ph.D. program, and who needs to break the news to one of them that he has not been admitted. In this case, the speaker's communicative intention is to update the applicant's beliefs about whether he'll be able to attend the professor's graduate program. The intention may also include a desired stance or approach to the goal, perhaps to deliver the news in a way that preserves the applicant's feelings as much as possible.

< Figure 1 about here >

The communicative intention inputs to a process of message encoding, which is tasked with determining the features of meaning termed the *message* that (subsequently) need to be linguistically encoded so that the speaker's communicative intention can be realized. Here, the message will include meaning (or referential) features that correspond to the addressee himself, the graduate program, and the status of being admitted to such a program (in this case in the negative), and the relations among these entities and states. The specific meaning features that are selected by message encoding processes can take into account the stance that the speaker wishes to take, and choose features of meaning that emphasize ego-neutral characterizations (e.g., by encoding a meaning that represents the applicant's fit to the program) rather than ego-threatening characterizations (e.g., by avoiding encoding a meaning that represents the applicant's poor qualifications).

Once the message is encoded, grammatical encoding processes can retrieve the relevant linguistic features that can express the features in the message. For example, grammatical encoding can retrieve the pronoun "you" to refer to the addressee, the words "our graduate program" to refer to the graduate program, and based on the meaning features selected to realize the intended approach, "not a good fit" to refer to the lack of admittance, and perhaps "I'm afraid that" to convey a sympathetic stance. Linguistic features also include the syntactic structures that order the selected words in an order that is permissible in English, and that relate all the entities appropriately to one another (e.g., structuring and ordering words so that the speaker conveys that the student hasn't been admitted to the program rather than the other way around). Once the set of linguistic features (words and structures) have been specified, they can be encoded into the features that will allow them to be spoken aloud -- phonological, phonetic, and prosodic features that ultimately determine how articulatory movements ought to be programmed to create the appropriate perceptible linguistic features.

Especially important here, following Bock (1982), the framework adopts the assumption that speakers can exert explicit or *executive control* at just two points in this process (as shown in Figure 1): First, executive control can be exerted upon the message-encoding process, such that speakers can with conscious effort determine which specific features of meaning ought to be linguistically encoded so as best to realize the speaker's communicative intention. Second, executive control can be exerted upon the final stages, where the acoustic features of the utterance are determined. Exerting executive control over message encoding is especially relevant to audience design, as it is at message encoding that a speaker can try to flexibly adapt which specific features of meaning are encoded into the message so as to accommodate some idiosyncratic feature of the communicative situation. Exerting executive control over phonetic encoding is also relevant to audience design, though in a more quantitative rather than qualitative way (e.g., by slowing or speeding phonological production).

Feedforward Audience Design

By itself, feedforward language production makes available two qualitatively distinct types of audience design. First, as mentioned above, the ability to exert executive control over message encoding and phonetic encoding allows a speaker to intentionally adjust the features of meaning that are to be linguistically encoded, or the way that an utterance is articulated, to accommodate specifics of the current communicative context. Second, presumably through learning, the process of grammatically encoding the message into a set of linguistic features may relatively automatically (i.e., without the involvement of executive control) encode meaning into linguistic features in such a way as to implement audience-design strategies. These tacit strategies that can be implemented during feedforward production are forms of what is here termed *feedforward audience design*. Importantly, feedforward audience design is heavily limited; in particular, contextual features that are unavailable to executive control mechanisms prior to the onset of utterance production cannot influence feedforward production and so drive an audience-design effect. And message features that trigger audience-design strategies cannot do so unless they are explicitly represented in the message prior to grammatical encoding beginning, and grammatical encoding strategies that implement audience design cannot take effect unless they have previously been learned (or are somehow emergent or innate).

A classic and well-known form of audience design that can be explained as resulting from the influence of executive control upon message encoding are *register effects*. For example, speakers speak differently to children than to adults, using shorter and more common words, and shorter sentences, each conveying less information (e.g., Snow, 1972). Additionally, such *child-directed speech* is usually spoken with exaggerated changes in pitch, and sometimes with higher baseline pitch. This modified way of speaking to children can be seen as a form of audience design; at least in terms of its ultimate effects on comprehension and learning, speakers' use of shorter words and sentences and more common words permit children to understand them despite the fact that children are (usually) less proficient. Exaggerated pitch changes and overall higher pitch is more likely to attract and maintain children's attention (which is often more limited). (For a review of issues relating to child-directed speech, see Bernstein Ratner, 2013.)

Within the current framework, the accommodations speakers make when addressing children can be seen as a form of feedforward audience design. The status of the addressee as a child, with the limited proficiency and attention span that that entails, is known to the speaker before the production of the utterance begins, and so is available to executive control mechanisms from the start so as to influence message and phonetic encoding processes. If the addressee's status as a child was not available to executive control mechanisms -- if the speaker were unaware that his or her addressee is a child -- then the speech would not take on the features of child-directed speech. Executive control uses the child status of the addressee to drive message encoding processes to represent simpler and more common meanings, for example, encoding a known pet as "dog" rather than "Pomeranian" (i.e., implementing a strategy of encoding at the more accessible basic level of conceptual representation rather than the less accessible subordinate level; Rosch, 1975). The child-directed register also can invoke a "diminutive" mode, whereby meaning features are encoded into the message that are more child-suitable, ultimately leading grammatical encoding to choose words such as “doggie” or “puppy.” Meanwhile, executive control over phonetic encoding can implement the register features including slower speech, higher pitch, and more exaggerated pitch changes.

A simpler form of audience design is the well-known *Lombard effect* (Lombard, 1911). People naturally increase their vocal effort when they cannot hear their own voice, for example, in a loud environment or while wearing headphones. This is a kind of audience-design effect, as increasing vocal effort increases the likelihood that speech will be audible to an addressee. The effect is involuntary, and is observed even in non-human animals (Potash, 1972). However, humans (at least) can control the effect, as they are able (with intent) to speak at a lower volume even when they are unable to hear their own voice. This shows that this is also feedforward audience design. The automatic nature of the effect itself suggests that within the framework, it arises from a direct influence of auditory perception mechanisms onto articulatory mechanisms, but the fact that it can be inhibited reveals a role for executive control over the accommodation (see also Buz et al., 2016).

A more subtle form of feedforward audience design was reported by Ferreira, Slevc, and Rogers (2005). Speakers were asked to describe sequences of line-drawings so that real versus hypothetical addressees could identify them in order. Critical drawings, when named in isolation, elicited *homophone* names, such as “bat” (depicting the flying mammal). Such critical pictures were named in displays where another picture was on some (*experimental*) trials the homophone counterpart (a baseball bat), but on other (*control*) trials was an object with an unrelated name (a goat). Thus, on experimental trials, if speakers named the critical picture with a bare homophone (“bat”), the description would be ambiguous, because it could apply also to the homophone counterpart; control trials had no such threat of ambiguity. The primary result that was reported was that in general, speakers were ineffective at avoiding ambiguity -- they only weakly avoided calling a flying mammal “bat” when it appeared in the same display as a baseball bat. But, in one experiment, speakers addressed either real addressees who were tasked with ordering the pictures based on speakers’ descriptions, versus hypothetical addressees (i.e., speakers were told, “describe the pictures so that an addressee could pick them out”). Results showed that speakers who addressed real addressees used fewer bare homophones across the board (on both experimental and control trials) than speakers who addressed hypothetical addressees. (See Kuhlen & Brennan, 2013, for a general account of how general audience properties, including the believability of the authenticity of an audience, can be a very important influence on the nature of audience design effects.)

Both of these results -- the failure to avoid bare homophone names when the homophone counterpart was in the display, but the overall greater avoidance of bare homophone names when addressing real addressees -- can be understood in terms of the feedforward production model shown in Figure 1. First, the greater overall avoidance of bare homophone labels with real addressees, like the effects described above, can be understood as a kind of register effect -- feedforward audience design. When real addressees were present, speakers (evidently) wanted to be better understood overall, and so produced more elaborate descriptions. For example, one speaker described the flying mammal with, “the animal bat with its wings open.” This is more information than an addressee needs; just “animal bat” would be enough to allow an addressee to distinguish the intended flying mammal from a baseball bat (for recent work on overspecification, see Heller et al., 2012; Koolen et al., 2011). But such overspecified utterances follow naturally if the presence of a real addressee compelled speakers, via executive control over message encoding, to linguistically encode an abundance of information so as to make the communication task especially likely to succeed.

Meanwhile, the failure to effectively avoid bare-homophone labels when targets were accompanied by their homophone counterparts falls neatly out of the principles embodied in the feedforward model. As noted, executive control can only influence the message-encoding process to accommodate a listener based on information available at the beginning of utterance planning. At the beginning of utterance planning, executive-control processes cannot represent that the critical object to be described potentially has the same name as another object in the display (in the experimental condition). Indeed, at the beginning of utterance planning, executive control does not even have available to it the name of the critical object at all -- at the lexical level, the point of feedforward production is to retrieve that object name. If executive control does not have the name of the critical object, it also won’t have the name of the homophone counterpart that causes ambiguity. In short, given the architecture shown in Figure 1, it is algorithmically impossible for feedforward production to avoid the type of ambiguity that arises when a to-be-described object has a name that is potentially confusable with another object in the same referential context -- the overlap of the two names is unavailable to the system at the time that it needs to decide what name to use for the critical object in the first place. (Note that the additional mechanisms described below that theoretically supplement feedforward production enable such ambiguity avoidance, however, when time and resources allow.)

The same type of explanation holds for speakers' inability to avoid garden-path sentences, described above. At the beginning of feedforward production, the system cannot know that it is at risk of formulating a sentence that includes a garden path. This is because "garden path" is a property of the linguistic utterance itself, and so that utterance must be formulated for the system to be able to diagnose that garden-path status.

Thus far, all audience design effects that have been described can be explained because a relevant feature (e.g., the child-status of an addressee) was available to production mechanisms before feedforward production began (or an effect should not be observed if the relevant feature is absent). This allows across-the-board message-encoding strategies to create a message that, when grammatically encoded, will lead to an utterance that is more likely to be comprehended. The other form of feedforward audience design (which is still part of feedforward production) can arise within the grammatical encoding process itself, having been "designed" into it by prior learning (or perhaps as an emergent property, or innately).

One (potential) audience-design effect that can be explained along these lines is the very heavily investigated *syntactic priming* effect (here more neutrally termed *syntactic repetition*). First reported experimentally by Levelt and Kelter (1982) and Bock (1986), syntactic repetition refers to the effect whereby a speaker will repeat recently experienced syntactic structures in their own subsequent speech. For example, a speaker who experiences a *prepositional dative* syntactic structure such as “The governess made a pot of tea for the princess” is likely to use another prepositional dative in a subsequent picture description, like “The boy is handing a paintbrush to the girl”; if instead the speaker had experienced a *double object dative* like, “The governess made the princess a pot of tea,” she or he will be more likely to use a double object such as “The boy is handing the girl a paintbrush” (for reviews, see Mahowald et al., 2016; Pickering & Ferreira, 2008). Although many of the early demonstrations of syntactic priming showed it arising when speakers merely read (and repeated) sentences purportedly as part of a memory task (e.g., Bock, 1986), other demonstrations have shown that speakers are syntactically primed by others’ speech (e.g., Branigan et al., 2000).

In an influential article, Pickering and Garrod (2004) argued that this type of syntactic repetition is an important form of audience design. To begin, Pickering and Garrod leveraged the observation that the repetition that is observed at different levels of linguistic (and hypothetically, non-linguistic) representation is non-independent. For example, if a speaker repeats the verb from a previously heard sentence in a subsequent picture description, even more repetition of syntactic structure is observed than if the speaker does not repeat the verb, an effect termed *the lexical boost* (Pickering & Branigan, 1998). Pickering and Garrod argued from the lexical boost that such non-independence may be pervasive, such that repetition of words and structures from one speaker to another should lead all the way up to the level of meaning (what they termed *situation models*; Zwaan & Radvansky, 1998); to “repeat” meaning from one conversational partner to another is one way to view the process of communication itself. In all, Pickering and Garrod argue that repetition of linguistic features among conversational partners is actually a reflection of an underlying *alignment* process that promotes efficient communication among those conversational partners.

Thus, within an alignment framework, syntactic repetition promotes more efficient communication and so can be viewed as a form of audience design. But syntactic repetition can be seen as communicatively relevant in terms that are more theory-neutral. If a conversational partner repeats syntactic structures at greater than chance rates, then that is face-valid evidence that that conversational partner heard and understood the structures that they are repeating.

As such, syntactic repetition can be seen as a basic form of audience design: A speaker who repeats structures she or he heard (in the default case, from another conversational partner) rather than producing an alternative is assisting the communicative process by doing so, either through alignment (Pickering & Garrod, 2004) or simply by signaling understanding (making syntactic repetition one of the many collateral signals in the sense of Clark, 1996). Considered as a form of audience design, different current explanations for syntactic repetition would frame it as a form of feedforward audience design. For example, Pickering and Branigan (1998) view syntactic repetition as following from the greater accessibility of syntactic knowledge from it having been recently experienced. This places the accommodation that syntactic repetition represents within the feedforward grammatical encoding process. Chang, Dell, and Bock (2006) explain syntactic repetition as a kind of learning effect -- that the original experience of hearing a type of event meaning expressed with a particular structure strengthens the connection between that event meaning and that structure, leading speakers to be more likely to use that structure again in the future. Jaeger and Snider (2013) attribute syntactic priming to learning with the goal to reduce "joint effort" of interlocutors and to "minimize the expected future prediction error" (for related discussion of the effects of repetition on prediction errors experienced during comprehension, see Fine & Jaeger, 2013, 2016; Fine et al., 2013; Myslin & Levy, 2016). Once again, the relevant strengthened connection sits within the feedforward grammatical encoding process, making it a form of feedforward audience design.

By itself, the repetition that comes from syntactic priming, to the extent that it is an audience-design effect, is best framed as an emergent property of the accessibility of recent structures (Pickering & Branigan, 1998) or learning processes (Chang et al., 2006; Dell & Chang, 2014; Fine & Jaeger, 2013; Jaeger & Snider, 2013). That is, nothing in these accounts describes the repetition as operating for audience-design purposes; the audience-design effects simply emerge "for free." However, learning is very likely implicated in a more striking result observed by Kaschak, Kutta, and Schatschneider (2011). They brought subjects into a particular lab setting, and led them to produce a large number of sentences with a particular structure (e.g., double-object dative structures). Then, a week later, subjects were brought to the same lab setting, and given the opportunity to produce either the previous structure (double object datives) or an alternative (prepositional datives). Results showed that speakers produced more double-objects than when they had been exposed to a large number of prepositional datives. In short, Kaschak et al. reported a syntactic repetition effect that lasted a week.

It is possible that the first week's experience in the Kaschak et al. (2011) experiment had an across-the-board effect on speakers, such that they used more double objects in all of their interactions at least leading up to the subsequent week's test. More likely, however, is that the repetition was conditioned on the lab setting -- that when producing language in the same setting, using about the same task, speakers' previous experience of having produced many of one type of structure led them to continue to produce the structure again. If so, then as a form of audience design, this is a combination of the two forms of feedforward audience design described thus far. The previous experience of producing many double-object structures in a particular lab setting (including all of the features of the lab and task) led speakers to learn of it as a kind of situating feature; when that situating feature is present, speakers produced a disproportionate number of a particular structure. Then, upon test, this situating feature is available to the feedforward production process before it begins, just like (for example) the child status of an addressee for register effects. This allows the grammatical encoding mechanisms that give rise to syntactic repetition (Chang et al, 2006; Pickering & Branigan, 1998) to condition the selection of the syntactic structure on the presence of the situating feature.

Three more studies illustrate how the availability of a communicatively relevant feature can combine with previously learned strategies to enact a feedforward audience design effect. Arguably the classic demonstration of audience design in production comes from Brennan and Clark (1996). They had speakers describe line-drawn objects in displays. One critical object, for example a shoe, was included among a number of other objects from other categories. Speakers described such objects to their addressees with their basic-level category name, that is, “shoe.” In subsequent rounds, displays included the shoe, but with other members of the shoe category as well -- the target loafer, along with a running shoe and a pump. Now, “shoe” isn’t sufficient, and so speakers changed their description to a subordinate level term such as “loafer.” In a final round, the same-category alternatives were removed, and speakers either spoke to the same addressee as in previous rounds, or a new addressee. With the same addressee, speakers continued to refer to the shoe as “loafer,” even though this level of specificity wasn’t needed. But to new addressees, speakers more often reverted to the basic-level “shoe.” That is, with a previous addressee, speakers maintained the historical precedent (see also Yoon & Brown-Schmidt, 2013), but with a new addressee, speakers represented that this historical precedent wasn’t in place, and so they referred to a preferred default form (see also Galati & Brennan, 2010). This effect can be seen as feedforward audience design: The status of the addressee as previous or new is available to the speaker before production is initiated, and so feedforward mechanisms can implement a strategy whereby historical precedents are maintained with previous addressees, but that such precedents are broken with new addressees (and see Lockridge & Brennan, 2003 for evidence that comprehenders are sensitive to whether speakers maintain or break such precedents).

Horton and Gerrig (2005; see also Horton & Gerrig, 2016) implemented a director-matcher paradigm, wherein a participant acting as director described pictures to be picked out by one of two participants (at different times across the task) acting as matchers. For half of directors, the types of pictures to be described aligned with matcher identity, such that directors described birds and dogs to one matcher but fish and frogs to the other. For the other half of directors, all types of pictures were described to both matchers. Then in a test phase, directors described all types of pictures to both matchers. Of interest is whether directors described pictures that were new to a particular matcher as if it were new, versus describing it as they had to the other matcher. The results showed that directors described pictures that were new to matchers appropriately as new when the picture types and matchers aligned than when they did not. In short, directors were better able to keep track of which pictures were new to which matchers when they could use the type of picture (dogs vs. fish) as a cue, which was only possible when matchers and picture types aligned. (See also Gorman et al., 2013, for relevant evidence.)

This effect can be understood along the lines of the explanation for the Kaschak et al. (2011) result. For directors who described picture types aligned with matchers, the type of picture (birds vs. dogs vs. fish vs. frogs) can act as a situating feature, such that directors can learn that some picture types (birds and dogs) should be associated with one matcher but the other picture types should be associated with the other. This can guide executive control mechanisms to cue message encoding to either rely on previously used labels (when the picture category cues that the picture is old for a matcher) or to specify meaning features more extensively (when the picture category cues that the picture is new). For directors whose picture types were not aligned with addressees, directors had to keep track of every individual picture-matcher combination, which overwhelms executive control mechanisms, and so feedforward audience design is not observed. (This is closely related to the idea that *one bit* distinctions can be used to support audience-design strategies; see Brennan et al., 2010).

The other result comes from Kurumada and Jaeger (2015). They looked at the production of Japanese, which is a *case-marking* language. This means that the roles of subject and object are indicated not by word order (as they are in English), but by suffixes or case-markers that affix to verbs. For example (using English words), a Japanese speaker would say something like “dog-ga cat-o chase” to say that “the dog chased the cat,” with “-ga” indicating the subject and “-o” the object. In casual spoken Japanese, however, the case markers -- though potentially useful -- are not always produced. Kurumada and Jaeger explored whether the choice to use or omit the object case marker (“-o”) follows audience-design principles.

In their Experiment 1, Kurumada and Jaeger (2015) determined whether the animacy of the object argument influences the use of case-markers. They had Japanese speakers say sentences that in English would be like “the teacher saw the student” versus “the teacher saw the fire engine.” When the object is animate (“student”), the sentence is potentially more confusable, because at least in principle, that object could be a subject (students can see teachers as much as teachers can see students). When the object is inanimate (“fire engine”), such confusability is largely absent (whereas teachers can see fire engines, fire engines typically cannot see teachers). Accordingly, Kurumada and Jaeger found that speakers were more likely to mention the object case marker when the object was animate and thus potentially confusable with the subject (see also Lee, 2006, for evidence from Korean). This is nicely explained as audience design, as speakers appear to be adding extra confusion-reducing content (the case marker) to their utterances just when their addressees are more likely to be confused.

Within the current framework, this can be explained as feedforward audience design. The animacy of the object argument is inherently represented within the message that inputs to the grammatical-encoding process. Through learning, grammatical encoding can develop the strategy that when it encodes a patient argument that is animate, it should be more likely to include the case-marking feature. (How this learning might come about is described below.) An analogous effect can be seen in English, whereby speakers are less likely to use a pronoun to refer to a person when more than one person in the described context has the same gender; instead, speakers are more likely to use a full noun phrase (Francik, 1985; Karmiloff-Smith, 1985). Again, the match or mismatch of the genders of arguments to be described is represented in the message prior to the start of grammatical encoding, so that grammatical encoding can develop a strategy of avoiding pronouns when referents' genders match (though this effect, like syntactic repetition, might be more of an emergent property, due to similarity based competition; see Gennari et al., 2012; Arnold & Griffin, 2007; MacDonald, 2013).

Summarizing, feedforward grammatical encoding, as illustrated in Figure 1, is capable of limited but effective forms of audience design, termed here feedforward audience design. In all cases, feedforward audience design is restricted to being contingent on information that is available to the feedforward production process *before* that process begins. What is central to all of these forms of audience design is that they do not require the speaker to make any inference about the particular knowledge state of the addressee, or of the potential communicative effect that a particular to-be-produced utterance might have with a particular addressee (see Dell & Brown, 1991, for a related proposal). Any such effect cannot be handled by feedforward production by itself. Instead, such audience design effects either should not be observed at all, or, if audience design contingent on the features of a particular utterance or a particular addressee is observed (as it has been -- see below), it will need to be explained as a form of recurrent processing audience design, described next. (A third possibility is, of course, that the present framework that organizes this review is incorrect in that it does not accurately describe how speakers engage in audience design.)

Recurrent Processing Audience Design: Empirical Effects

The core of feedforward audience design is that it stems from the feedforward production process -- what is sometimes called “one shot” production. Language users can, of course, be much more deliberative than this. Given time and resources, we are able to go beyond setting up a message that can grossly accommodate addressees based on previously learned strategies (i.e., feedforward audience design); we are sometimes able to evaluate the potential communicative effect that a particular utterance might have for a particular addressee. Within the current framework, this is termed *recurrent processing audience design*. Before describing a mechanism for recurrent processing audience design, this review describes relevant effects to be explained.

The rough idea behind recurrent processing audience design is that given time and resources, a speaker can formulate a candidate utterance and “try it out” on a particular addressee in a particular situation. The signature result illustrating this type of audience design is Horton and Keysar (1996). Speakers and addressees viewed separate parts of a display that showed pairs of simple shapes. On each trial, a *target* shape moved from the speaker's to the addressee's side of the display; sometimes this shape changed en route (though participants did not know when it did or did not change). The speaker's task was to describe the target shape so that the addressee could know if it changed. The target shape invited a contrast to another *context* shape that was also in the display. Then, half of subjects were told that their addressee could see the same context shape; the other half were told that their addressees had no context shape. The question of interest is whether speakers would describe the target via contrast with the context shape (e.g., by saying "larger circle"), even when speakers knew addressees didn't have a context shape. Results showed that with time pressure, speakers relatively frequently described targets in contrast with the context shape, even when told their addressees couldn't see the context. But without time pressure, this effect attenuated. In sum, given time and resources, speakers were able to take into account that their addressees had a different perspective than they did; without time and resources, they were less able to do so.

Above, one experiment from Kurumada and Jaeger (2015) was described, showing that Japanese speakers mentioned an optional object case-marker more when the object was animate than when it was inanimate, which was argued to be a form of feedforward audience design. Two other experiments in the same paper reported a more subtle form of audience design. In these, speakers produced sentences in which the relationships within the described event were either plausible or implausible. For example, speakers were led to produce either that a police officer arrested a criminal (plausible) or that a criminal arrested a police officer (implausible). Implausible sentences are likely to be misinterpreted by an addressee in favor of its more plausible counterpart (e.g., F. Ferreira, 2003). If speakers wish to avoid such misinterpretation, they should include the optional object case-markers more when the event being described is implausible rather than plausible. This particularly clever form of audience design is indeed what Kurumada and Jaeger observed.

The difference between the two sorts of results from Kurumada and Jaeger (2015) can be used to illustrate the key difference between feedforward and recurrent processing audience design. Audience design contingent on the animacy of the object argument (described above) is compatible with feedforward audience design, because the animacy of the object argument is available and represented within the preverbal message before grammatical encoding begins. Thus, feedforward production can increase the probability of using the object case-marker, assuming such a grammatical-encoding strategy is in place (presumably having already been learned -- see below). But the effect of implausibility on object-case marking is more complex. All elements (police officers, criminals, arresting) are indeed represented in the message before grammatical encoding starts (at least in principle). But the *plausibility* of the relationship among those elements is presumably *not* represented within the preverbal message. The relative plausibility of an event is complex and multidimensional, and contingent on all kinds of real-world knowledge. Comprehension mechanisms -- which are tasked with understanding an utterance and fitting the resulting meaning into what the comprehender knows about the real world -- are well suited for assessing plausibility; production mechanisms, which already “know” speakers’ intended meanings before feedforward production begins, are not. As such, this effect of (im)plausibility is extremely *un*likely to be feedforward audience design. Here, it is framed as recurrent processing audience design.

All of the results described in this section so far describe audience-design effects where, within the current framework, speakers can "try out" a candidate utterance so as to potentially modify it for the benefit of addressees' comprehension. Indeed, many current theories of language production include just such a process, called a *monitor* (Postma, 2000), and in particular, a *perceptual loop monitor* (Levelt, 1989; Hartsuiker & Kolk, 2001). The idea behind the perceptual loop monitor is that the comprehension system can be enlisted to comprehend some product of the feedforward production process. The result of that comprehension process can then be used to potentially modify the production plan, if that result indicates that the speaker's communicative intention may not be fulfilled with the already formulated utterance.

According to many accounts, however, relying upon the production monitor can be costly (see Hartsuiker & Kolk, 2001). If the production process customarily computes candidate utterances in order to evaluate their potential for communicative success, this would seem to introduce significant inefficiency into the production process -- and the more comprehensively the system pursues this strategy, the greater the inefficiency. For example, in principle, a system of this type could aim to ensure that the first words of an utterance are maximally communicatively effective; if so, then even the initiation of production needs to be suspended until the production process formulates the beginning of the utterance, the comprehension system comprehends it, and an evaluation process assesses whether comprehension is likely to be as intended. Considerations such as these have led to a range of proposals that production is “egocentric” (Wardlow Lane & Ferreira, 2010), or that speakers formulate candidate utterances egocentrically and adjust them only if resources allow (Horton & Keysar, 1996).

A way through these problems may come, however, from introducing principles from *motor control theory* into the production account (as language production is, in the limit, a complex act of motor control). In particular, the idea of *forward models* has proven valuable in many accounts of motor behavior (Grush, 2004; Wolpert et al., 2003), including as explanations for the problem of how the production system learns to produce speech sounds (see Tourville et al., 2014). In the higher-level domain of audience design, however, the use of forward models is more complex (Wolpert et al., 2003). Even so, general forward modeling principles may indeed be an informative way to allow more complex recurrent processing audience design effects without the heavy cost in terms of time and effort that comes from using a full-on perceptual loop monitor.

Recurrent Processing Audience Design: A Forward-Modeling Approach

In order to circumvent the costs in terms of time and resources that come from using the perceptual-loop monitor to explain recurrent processing audience design, other costs, in terms of a more complex cognitive architecture, must be incurred. Such mechanistic costs come primarily from postulating a separate, independent mechanism for anticipating the communicative consequences of candidate utterances. Some specifics regarding the nature of this mechanism are inspired by motor control theory, and in particular, forward models. This proposal expands on ideas described in Jaeger and Ferreira (2013), Kurumada and Jaeger (2015) and Buz et al. (2016).

A forward model is a separate mechanism that has the task of simulating or emulating (Grush, 2004) the consequences of some behavioral choices before those behavioral choices are finalized. Such a mechanism enables two key benefits. First, by stripping it down to the minimum needed to be able to predict relevant consequences, the forward modeling process can operate more quickly than the mechanisms driving the actual behaviors (and therefore, given its nature, more quickly than perceptual loop monitoring). In turn, this allows the match of the predicted effects with the intended effects to act as a corrective signal for those mechanisms that drive the actual behaviors. Second, by predicting communicative consequences before they happen, the system can compare the predicted consequences against the actual consequences; if they match, then the system can continue using behavioral and forward-modeling mechanisms whose veracity has been confirmed (because the predicted and actual consequences matched). If predicted and actual consequences do not match, then the system can use the difference between predicted and actual outcome as an error signal, to correct either the cognitive mechanisms that drive the overt behavior, or the forward modeling mechanisms that generated the predictions (a blame-assignment mechanism must intervene to estimate which mechanism(s) ought to undergo learning).

An oft-used analogy for describing forward modeling comes from ocean navigation (e.g., Grush, 2004). Imagine a ship's crew that wishes to navigate to a goal location. The crew can position sails and rudders so as to direct the ship toward the goal. But rather than just position the sails and rudders and assume success (which would be analogous to using feedforward production alone), the crew can continue to monitor and modify the navigational settings en route. In particular, using formal navigational principles, the ship's crew can predict the observed outcomes in terms of the positions of stars, fixed landmarks, and so forth that ought to result from the known positions of the sails, rudder, wind speed, direction of the current, and so forth, at some given time point in the future. Once that future time point arrives, the crew can compare the predicted positions of the stars, landmarks, and so forth to the actual positions of the stars, landmarks, and so forth. If prediction and observation match, all is good, and the ship can continue on to its goal. If prediction and observation do not match, then something must be corrected. Either the crew erred in generating predictions, or the anticipated navigational consequences of the sail positions, rudder position, etc., were different from reality. If the crew can be confident of their predictions, then they can use the disparity between the predicted outcome and the actual outcome to change the positions of the sails, rudder, etc., to correct (compensating for the unknown factor that led to the discrepancy), so as to continue toward the goal position.

The key components of this analogy are that a copy of the system's intent (the ship's desired location) and plan (the configuration of the navigational systems) are sent to a forward-modeling process (the ship's crew, which use their own knowledge, tools such as maps and sextants etc.). This forward model is tasked with generating predicted outcomes. The predicted outcomes are compared to actual outcomes to (in the typical case) correct the behavior of the system before that behavior is complete, if such correction is needed.

As noted above, this type of process has already been informatively applied to low-level language processes, in particular, for the systems that map syllable plans onto articulatory gestures (Tourville et al., 2014). For audience design, the problem is more complex (and more abstract), but the general principles are the same. A sketch of an architecture is illustrated in Figure 2. The same communicative intent (analogous to the goal location) and message (analogous to the position of the sails and rudders, etc.) that is input to the feedforward production process is sent also to the forward modeling process. The forward modeling process generates a set of communicatively relevant features that it estimates will be generated by the feedforward production process. These communicative features can be input to an evaluation process that represents the particulars of the addressee (analogous to wind speed, direction of the currents, etc. -- idiosyncratic features of the context), and uses these together to generate predictions of the communicative effect of the generated communicative features. Indeed, the evaluation process can simply be the language-comprehension process, which itself has the function of inferring meaning and ultimately the communicative effect that would derive from a comprehended utterance. If the predicted communicative effect of the forward model's generated linguistic features matches the speaker's actual intent, the feedforward production process can proceed based on the original message -- the feedforward process should lead to behaviors that will comport with the speaker's intention. But if the predicted communicative effect mismatches the speaker's intent, the discrepancy can be provided as input to the message encoding process, which can modify the message so as to make the communicative intent more likely to be realized.

< Figure 2 about here >

For example, imagine again the professor who intends to inform the Ph.D. applicant that he will not be admitted to her graduate program. If she is not yet experienced at gently delivering negative news (i.e., her message and grammatical encoding processes have not yet learned how transform a sympathetic stance that is represented in the communicative intention through the message into words like “not a good fit”), her feedforward production process and -- importantly for present purposes -- her forward model process will instead first generate words such as, “scores are too low.” These words can then be input to a context-specific comprehension-evaluation process, one that is better equipped to take the perspective of the Ph.D. applicant. The comprehension process can predict that the communicative effect of “scores are too low” will be hurt feelings, which diverges from the professor’s communicative intent to be sympathetic. That difference can then be fed back to the message encoding process, to guide it to choose a different set of words, perhaps nudging in a direction that is more distant from the ego of the Ph.D. applicant, and so less likely to result in hurt feelings.

This type of forward-modeling mechanism can explain the recurrent processing audience design effects described above. An illustrative example is Horton and Keysar (1996). Recall that they showed that without time pressure, speakers were able to adopt their addressees’ divergent perspectives, and so avoid describing a shape as “small circle” when there was a larger comparison circle in the speakers’ view, but not in the addressees’ view. With time pressure, however, speakers more often referred (via comparison) to the privileged context object. Within this framework, the audience design without time pressure operates via the full forward-modeling mechanism: With time and resources, a speaker can use a forward modeling process to generate the communicatively relevant feature (the comparison adjective), and evaluate that its communicative effect for the addressee will be inappropriate (because the evaluation-comprehension process -- but not the feedforward production process -- represents that the addressee cannot see the context object). This divergence from the speaker's communicative intent (to refer appropriately) can be fed back to message encoding, along with a signal that the problem lies with the adjective. With time pressure, this full process plays out less successfully, and so the speaker refers to the privileged context object more.

The results of Experiments 2 and 3 in Kurumada and Jaeger (2015) can be explained similarly. Recall that these experiments showed that implausible relationships ("the criminal arrested the cop") resulted in more case-marking than plausible relationships ("the cop arrested the criminal"). This effect of plausibility is likely to be recurrent processing audience design -- that the intent and message are sent to a forward-modeling process that computes the communicatively relevant features -- here, minimally, the two nouns and (if present) their case markers, and the verb. These features can be fed to the evaluation-comprehension process that is well equipped to estimate plausibility. If the nature of the agent-patient relationship (as understood by the comprehension process) along with the absence of an object case-marker is deemed by the evaluation process to be likely to be misunderstood, then it will lead to a discrepancy -- an error signal -- between the speaker's intent and the predicted communicative effect of the generated linguistic features. In turn, this error signal can be fed back to message encoding so that the message can be modified, presumably so as to more clearly mark the agent-patient relationship, which in turn can accomplished by formulating the object case-marker.

It is important to note that the forward model *generates* communicatively relevant linguistic features for the purpose of *predicting* a communicative effect. The forward model does not predict the linguistic features -- the linguistic features are the devices that are used to generate the predicted communicative effect. Other approaches that use forward modeling in production (Pickering & Garrod, 2013) claim to predict linguistic features, and are silent with respect to the prediction of communicative effects; see Jaeger and Ferreira (2013) for further discussion.

It is legitimate to ask why the system should employ an independent forward-modeling process to generate the communicatively relevant features -- a kind of duplicate processing mechanism -- rather than use the actually formulated linguistic features generated by the feedforward production process. Two related factors favor postulating an independent process. First, feedforward production must compute *all* linguistic features needed for actual articulation, whereas the forward-modeling process need only compute communicatively relevant features. For example, for the sake of illustration, assume that in English, subject-verb agreement is only weakly communicatively relevant -- that an addressee will infer about the same meaning and intent for “the cat meow” as for “the cat meows.” If so, then the forward-modeling process need not compute subject-verb agreement. Feedforward production, however, still must, as native speakers still generally get subject-verb agreement right.

The second (related) reason is speed: Because the forward-model process need not formulate all of the linguistic detail that feedforward production needs to formulate, it can do so more quickly. This provides more time for the evaluation-comprehension process to estimate the communicative effect of the generated linguistic features, compared to if the products of the feedforward production process were used instead (which would reduce the approach to a standard perceptual loop; see Postma, 2000).

How does feedforward production obtain audience-design strategies?

The forward-modeling process that is claimed here to underlie recurrent processing audience design yields an additional benefit: Because it predicts communicative effects, by comparing intended effects to predicted effects (or predicted effects to actual effects), it can use any resulting error signal to tune production strategies so that future production is more likely to be communicatively successful. For example, consider one last time the professor who needs to inform the applicant that he was not admitted to her Ph.D. program. The error signal (derived by comparing the predicted communicative effect from the forward-model-generated linguistic features against the intended effects) that indicates that “scores are too low” will result in hurt feelings can not only signal to message-encoding mechanisms that the message needs to be tweaked to satisfy the speaker’s communicative intent, it can also be used to modify grammatical encoding strategies so that in the future, when a sympathetic stance is represented as part of the speaker’s communicative intention, more ego-neutral meanings can be encoded into speaker’s message the first time around. That is, the first time the professor intends to express unwelcome news sympathetically, she may need to engage in more effortful recurrent processing audience design to predict the communicative consequences of a set of generated linguistic features. But if the predicted communicative effects are used to train relevant aspects of the feedforward production process itself, then subsequently, recurrent processing audience design will no longer be needed. Instead, message encoding strategies can follow a feedforward audience-design strategy, and map the relevant aspect of the communicative intention (the sympathetic stance) to the message ("not a good fit"). An architecture sketching out all audience design components is shown in Figure 3.

< Figure 3 about here >

Relevant experimental evidence comes from Buz et al. (2016; see also Seyfarth, Buz, & Jaeger, 2016). Directors and (simulated) matchers participated in a simple task over the internet where they saw three printed words, and directors named one of those words so that the matchers could pick it out. Buz et al. manipulated two factors. First, for a given target (e.g., "pig"), one distractor was sometimes a close competitor (e.g., "big") or it was not (e.g., "heart"). Crossed with this, directors received either no feedback as to addressees' success, received only positive feedback (i.e., were signaled that addressees were always correct), or received positive and negative feedback (i.e., were signaled that addressees were sometimes sometimes incorrect, because the addressee chose the close competitor). Two results of interest were observed. First, directors articulated targets such that they were more distant from close competitors, but only when those close competitors were present (e.g., directors pronounced "pig" as more different from "big," but only when "big" was in the display rather than "heart"). This effect was observed both when directors received no feedback and when they received only positive feedback. But when directors received negative (as well as positive) feedback indicating that matchers sometimes chose the close competitor, the tendency to articulate targets to be different from close competitors was even greater than when directors received no or only positive feedback.

The primary tendency to articulate “pig” as more different from “big,” but only when “big” is present, is likely to at least initially be a form of recurrent processing audience design. In their task, Buz asked directors on each trial to silently read the three candidate words prior to being informed of the target; this experience informs directors’ model of their addressee such that if “pig” is articulated it could sound very similar to “big,” which would threaten the directors’ intention to say targets so that they are accurately picked out. This discrepancy between directors’ predicted outcome (that the addressee might pick “big”) and desired outcome (that they want the addressee to pick “pig”) can be used by executive control mechanisms to influence the articulation process to hyperarticulate targets, but only when the close competitor was present.

After some experience with this task context, however, the hyperarticulation strategy could become a form of feedforward audience design. Because Buz et al used printed words, the presence of a close competitor was cued by the similarity in spelling between the target (“pig”) and the competitor (“big”). That similarity in spelling is available to feedforward production before production begins. And so via the learning mechanism described above, feedforward production can learn a new feedforward audience-design strategy of hyperarticulating targets when a visually similar distractor is present. This is in fact how Buz and colleagues explain their results. (Indeed, the present framework makes an interesting prediction: If a task like Buz et al.’s is used but with pictures as stimuli instead of words [and stimuli don’t repeat, and subjects are prevented from silently naming stimuli before overt naming], little or no hyperarticulation should be observed, because like Ferreira et al., 2005, the potential phonological similarity between the target and competitor will not be available to feed forward production at the beginning of the naming process.)

The effect whereby directors hyperarticulated targets even more when receiving negative feedback is obviously a learning effect. Within the current framework, it can be accommodated two ways: explicitly or implicitly. Explicitly, directors might consciously recognize the implications of the negative feedback (that “big” is confusable with “pig”), and then via executive control, direct articulatory mechanisms to hyperarticulate targets. They can then learn this as an explicit strategy, in the same way that speakers explicitly learn to refer to an acquaintance by first and last name when they know that the first name is shared by more than one person. More interesting in the current context is the implicit possibility. In particular, the perceived communicative failure (i.e., the negative feedback) can act as an error signal to tune a feedforward audience-design strategy, so that the presence of visually similar candidate words on a future trial triggers greater hyperarticulation (Buz et al. present evidence pointing toward the implicit rather than explicit possibility).

Another report of speakers learning an audience design effect that can fit within the current framework comes from Fedzechkina, Jaeger, and Newport (2012; see also Fedzechkina, Newport, & Jaeger, 2017). They had native speakers of English learn an artificial language that (like Japanese) sometimes used case marking suffixes to indicate the object grammatical role. Like Experiment 1 of Kurumada et al. (2015), object arguments were sometimes animate and sometimes inanimate, and in the language speakers were exposed to, half the time an object case marker (when it appeared) was used with an animate direct object and half the time with an inanimate direct object. Interestingly, as speakers came to learn the artificial language and use it themselves, then like the subjects in the Kurumada and Jaeger experiment, they tended to use case markers more with animate than with inanimate object arguments. Like Kurumada and Jaeger, this can be argued to be an audience-design effect, because animate object arguments have more potential to be confused with subject arguments, and any such confusion will be reduced by explicitly marking the object arguments.

This effect follows naturally from the learning mechanism described above. As speakers came to learn and use the artificial language, recurrent processing audience design mechanisms can formulate candidate utterances and evaluate their potential for communicative success (which implicitly at least is the speaker's communicative intention). When forward-modeling mechanisms generate candidate linguistic features that include animate-object arguments without an object case marker, evaluation mechanisms can predict that subject and object arguments are more confusable and so communicative failure is more likely (see Fedzechkina, 2012, for a related argument). This discrepancy between the speaker's communicative intention and the predicted communicative effect can act as an error signal that can be used to train grammatical encoding (and forward modeling) mechanisms to make it less likely in future formulations that animate object arguments are formulated without case markers. In contrast, when an inanimate object argument is formulated without a case marker, less potential for confusion is predicted (because of the asymmetry in animacy between subject and object), and so the same error signal is absent (and depending on the learning strategy that's involved, the tendency to exclude the object case marker with the inanimate object argument might even be strengthened).

General predictions

Throughout, various predictions that follow from this framework have been described. Here, we collect seven predictions that fall directly out of the described framework:

(1) Feedforward audience-design strategies can only act on features that are available before the act of production begins. For example, the animacy of an argument is available before production begins, and so can be the basis of a feedforward audience design strategy (e.g., including optional case markers; Kurumada & Jaeger, 2015, Experiment 1). Plausibility (Kurumada & Jaeger, 2015, Experiments 2 and 3) is unlikely to be available to feedforward production mechanisms before production begins, and so cannot be the basis of a feedforward audience design strategy.

(2) This is a special case of Prediction (1), but is important enough to mention separately: Any form of audience design that's contingent on properties of the produced utterance itself cannot come from feedforward audience-design mechanisms. Feedforward audience design comes from feedforward production mechanisms; until feedforward production has operated, properties of the produced utterance are unavailable.

(3) When an audience-design strategy is contingent on an idiosyncratic property of the addressee (e.g., an atypical knowledge state of this particular addressee) or of the situation, recurrent processing audience-design mechanisms are likely to be needed. This is because feedforward production can only use strategies that it has already learned or that (via executive control) are "common sense."

Idiosyncratic properties of addressees and contexts are unlikely to have been previously learned or follow from "common sense."

(4) Recurrent processing mechanisms should *not* implement audience-design strategies based on communicatively irrelevant features of an utterance. An interesting research project would be to first test whether a feature such as grammatical agreement is communicatively relevant (e.g., whether comprehenders customarily disambiguate a high- or low-attached relative clause based on grammatical agreement). If not, then even when recurrent processing audience design mechanisms are involved, speakers should be unable to determine that an utterance will not fulfill their communicative intention based on the agreement feature.

(5) The mechanisms that underlie recurrent processing audience-design effects are more resource sensitive than the mechanisms that underlie feedforward audience-design effects, as the former require the time and effort for the features generated by the forward model to be evaluated for their communicative effects, with that evaluation fed back to message-encoding processes to tweak the formulation process. This provides a litmus test for a claim that a particular effect is a recurrent-processing audience-design effect: If so, then if cognitive pressure is applied (e.g., time pressure, concurrent verbal load, etc.), the degree to which the audience design is observed should diminish. Note that together with Prediction (1) above, this provides a means of falsifying the current framework: If a particular feature upon which an audience-design strategy is based is unlikely to be available to feedforward mechanisms before production begins (e.g.,

plausibility), but if cognitive pressure does not diminish the effect, then this framework cannot explain the effect.

(6) In contrast, the mechanisms that underlie feedforward audience-design effects are less resource sensitive (see Bock, 1982). Thus, a litmus test for the claim that a particular effect is feedforward audience-design is that it should be relatively insensitive to cognitive pressure manipulations.

(7) Feedforward audience design strategies should not exist if they cannot be learned, they are not otherwise motivated by the way production mechanisms work (i.e., are "emergent" in the terminology above), or are innate.

Together, the assumptions sketched out in the architecture in Figure 3, along with the set of principles described here, provides a quite constraining framework for understanding audience-design effects, even if the framework is complex. (For computational-level analyses that seem consistent with this framework, see Frank, 2017; Heller, Parisien, & Stevenson, 2016.)

Summary and conclusions

Speakers speak to be understood. As such, they should endeavor to craft their utterances so that those utterances are as easy to understand as possible. This however is no easy task. The framework outlined above aims to situate a range of audience-design effects in a set of reasonable assumptions about the way that production mechanisms are likely to operate. Whether the framework is borne out by future data remains to be seen, but in any case, it should help us to better understand what speakers do and do not do -- and importantly, why -- to make themselves better understood to their addressees.

References

Arnold, J., & Griffin, Z. M. (2007). The effect of additional characters on choice of referring expression: Everyone counts. *Journal of Memory and Language*, *56*(4), 521–536.

Arnold, J. E., Wasow, T., Asudeh, A., & Alrenga, P. (2004). Avoiding attachment ambiguities: The role of constituent ordering. *Journal of Memory & Language, 51*, 55-70.

Bernstein Ratner, N. (2013). Why talk with children matters: Clinical implications of infant-and child-directed speech research. In *Seminars in speech and language* (Vol. 34, pp. 203-214): Thieme Medical Publishers.

Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review*, *89*, 1–47.

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18*, 355-387.

Bock, J. K., & Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945-984). San Diego: Academic Press.

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition, 75*(2), B13-B25.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *22*(6), 1482–1493.

Brennan, S. E., Galati, A., & Kuhlen, A. K. (2010). Two Minds, One Dialog: Coordinating Speaking and Understanding. In B. H. Ross (Ed.), *Psychology of Learning and Motivation* (Vol. 53, pp. 301–344). Academic Press.

Buz, E., Tanenhaus, M. K., & Jaeger, T. F. (2016). Dynamically adapted context-specific hyper-articulation: Feedback from interlocutors affects speakers' subsequent pronunciations. *Journal of Memory and Language, 89*, 68-86. doi:10.1016/j.jml.2015.12.009

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review, 113*(2), 234-272.

Clark, H. H. (1996). *Using language*. Cambridge, England: Cambridge University Press.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93*, 283-321.

Dell, G. S., & Chang, F. (2014). The P-chain: relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B-Biological Sciences, 369*(1634). doi:ARTN 20120394 10.1098/rstb.2012.0394

Elsness, J. (1984). *That* or zero? A look at the choice of object clause connective in a corpus of American English. *English Studies., 65*, 519-533.

Fedzechkina, M., Jaeger, T. F., & Newport, E. L. (2012). Language learners restructure their input to facilitate efficient communication. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(44), 17897–17902.

Fedzechkina, M., Newport, E. L., & Jaeger, T. F. (2017). Balancing Effort and Information Transmission During Language Acquisition: Evidence From Word Order and Case Marking. *Cognitive Science*, *41*(2), 416–446.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, *47*(2), 164–203.

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language*, 25, 348-368.

Ferreira, V. S. (2008). Ambiguity, accessibility, and a division of labor for communicative success. *Psychology of Learning and Motivation: Advances in Research and Theory, 49*, 209-246.

Ferreira, V. S., & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology, 40*, 296-340.

Ferreira, V. S., & Hudson, M. (2011). Saying "that" in dialogue: The influence of accessibility and social factors on syntactic production. *Language and Cognitive Processes, 26*, 1736-1762.

Ferreira, V. S., & Schotter, E. R. (2013). Do verb bias effects on sentence production reflect sensitivity to comprehension or production factors? *Quarterly Journal of Experimental Psychology, 66*(8), 1548-1571. doi:10.1080/17470218.2012.753924

Ferreira, V. S., & Slevc, L. R. (2007). Grammatical encoding. In M. G. Gaskell (Ed.), *The Oxford Handbook of Psycholinguistics* (pp. 453-469). Oxford: Oxford University Press.

Ferreira, V. S., Slevc, L. R., & Rogers, E. S. (2005). How do speakers avoid ambiguous linguistic expressions? *Cognition, 96*, 263-284.

Fine, A. B., & Jaeger, T. F. (2013). Evidence for Implicit Learning in Syntactic Comprehension. *Cognitive Science, 37*(3), 578-591. doi:10.1111/cogs.12022Fine, A. B., Jaeger, T. F., Farmer, T. A., & Qian, T. (2013). Rapid Expectation Adaptation during Syntactic Comprehension. *Plos One, 8*(10). doi:ARTN e77661 10.1371/journal.pone.0077661

Francik, E. (1985). Referential choice and focus of attention in narrative. *Unpublished Doctoral Dissertation, Stanford University*.

Frank, M. C. (2017). Rational speech act models of pragmatic reasoning in reference games. *PsyArXiv*. http://doi.org/10.17605/OSF.IO/F9Y6B

Fukumura, K. (2018). Ordering adjectives in referential communication. *Journal of Memory and Language*, *101*, 37–50.

Galati, A., & Brennan, S. E. (2010). Attenuating information in spoken communication: For the speaker, or for the addressee? *Journal of Memory and Language*, *62*(1), 35–51.

Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 9, pp. 133-177). New York: Academic Press.

Gennari, S. P., Mirković, J., & Macdonald, M. C. (2012). Animacy and competition in relative clause production: a cross-linguistic investigation. *Cognitive Psychology*, *65*(2), 141–176.

Gorman, K. S., Gegg-Harrison, W., Marsh, C. R., & Tanenhaus, M. K. (2013). What's learned together stays together: speakers' choice of referring expression reflects shared experience. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *39*(3), 843–853.

Grush, R. (2004). The emulation theory of representation: motor control, imagery, and perception. *The Behavioral and Brain Sciences*, *27*(3), 377–396.

Hartsuiker, R. J., & Kolk, H. H. J. (2001). Error monitoring in speech production: A computational test of the perceptual loop theory. *Cognitive Psychology*, *42*(2), 113–157.

Heller, D., Parisien, C., & Stevenson, S. (2016). Perspective-taking behavior as the probabilistic weighing of multiple domains. *Cognition*, *149*, 104–120.

Horton, W. S., & Gerrig, R. S. (2005). The impact of memory demands on audience design during language production. *Cognition*, *96*, 127–142.

Horton, W. S., & Gerrig, R. J. (2016). Revisiting the Memory-Based Processing Approach to Common Ground. *Topics in Cognitive Science*, *8*(4), 780–795.

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, *59*(1), 91–117.

Jaeger, T. F. (2006). *Redundancy and syntactic reduction in spontaneous speech. Unpublished doctoral dissertation*. Stanford University, Stanford, CA.

Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology, 61*, 23-62. doi:Doi 10.1016/J.Cogpsych.2010.02.002

Jaeger, T. F. (2011). Corpus-based research on language production: Information density and reducible subject relatives. In E. M. Bender & J. E. Arnold (Eds.), *Language from a Cognitive Perspective: Grammar, Usage, and Processing. Studies in honor of Tom Wasow* (pp. 161-197). Stanford, CA: CSLI Publications.

Jaeger, T. F., & Ferreira, V. (2013). Seeking predictions from a predictive framework. *The Behavioral and Brain Sciences*, *36*(4), 359–360.

Jaeger, T. F., & Snider, N. E. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition, 127*(1), 57-83. doi:10.1016/j.cognition.2012.10.013

Karmiloff-Smith, A. (1985). Language and cognitive processes from a developmental perspective. *Language and Cognitive Processes*, *1*, 61–85.

Kaschak, M. P., Kutta, T. J., & Schatschneider, C. (2011). Long-term cumulative structural priming persists for (at least) one week. *Memory & Cognition*, *39*, 381–388.

Kuhlen, A. K., & Brennan, S. E. (2013). Language in dialogue: When confederates might be hazardous to your data. *Psychonomic Bulletin & Review, 20*(1), 54-72. doi:10.3758/s13423-012-0341-8

Kurumada, C., & Jaeger, T. F. (2015). Communicative efficiency in language production: Optional case-marking in Japanese. *Journal of Memory and Language*, *83*, 152–178.

Lee, H. (2006). Parallel optimization in case systems: Evidence from case ellipsis in Korean. *Journal of East Asian Linguistics*, *15*(1), 69–96.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levelt, W. J. M., & Kelter, S. (1982). Surface form and memory in question answering. *Cognitive Psychology, 14*, 78-106.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral & Brain Sciences, 22*(1), 1-75.

Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review, 98*, 122-142.

Lockridge, C. B., & Brennan, S. E. (2002). Addressees' needs influence speakers' early syntactic choices. *Psychonomic Bulletin & Review*, *9*(3), 550–557.

Lombard, E. (1911). Le signe de l'elevation de la voix. *Annales des maladies de l'oreille, du larynx du nez et du pharynx, 37*, 101-119.

Macdonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology*, *4*, 226.

Mahowald, K., James, A., Futrell, R., & Gibson, E. (2016). A meta-analysis of syntactic priming in language production. *Journal of Memory and Language, 91*, 5-27. doi:10.1016/j.jml.2016.03.009

Myslin, M., & Levy, R. (2016). Comprehension priming as rational expectation for repetition: Evidence from syntactic processing. *Cognition, 147*, 29-56. doi:10.1016/j.cognition.2015.10.021

Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory & Language, 39*(4), 633-651.

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin, 134*, 427-459.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences, 27*(2), 169-226.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *The Behavioral and Brain Sciences*, *36*(4), 329–347.

Postma, A. (2000). Detection of errors during speech production: A review of speech monitoring models. *Cognition*, *77*, 97–131.

Potash, L. M. (1972). Noise-induced changes in calls of Japanese-quail. *Psychonomic Science, 26*(5), 252-254.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General, 104*, 192-233.

Seyfarth, S., Buz, E., & Jaeger, T. F. (2016). Dynamic hyperarticulation of coda voicing contrasts. *The Journal of the Acoustical Society of America*, *139*(2), EL31–EL37.

Snow, C. E. (1972). Mothers' speech to children learning language. *Child Development, 43*, 549-565.

Temperley, D. (2003). Ambiguity avoidance in English relative clauses. *Language, 79*, 464-484.

Tourville, J. A., Peeva, M. G., & Guenther, F. H. (2014). Perception-Production Interactions and their Neural Bases. In Goldrick, M. Ferreira, V. S. Miozzo, M.

(Ed.), *Oxford Handbook of Language Production* (pp. 460–477). Oxford University Press.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. Journal of Memory and Language, 33, 285-318.

Wardlow Lane, L., & Ferreira, V. S. (2008). Speaker-external versus speaker-internal forces on utterance form: Do cognitive demands override threats to referential success? *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *34*, 1466–1481.

Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *358*(1431), 593–602.

Yoon, S. O., & Brown-Schmidt, S. (2013). Lexical differentiation in language production and comprehension. *Journal of Memory and Language*, *69*(3), 397–416.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*, 162-185.

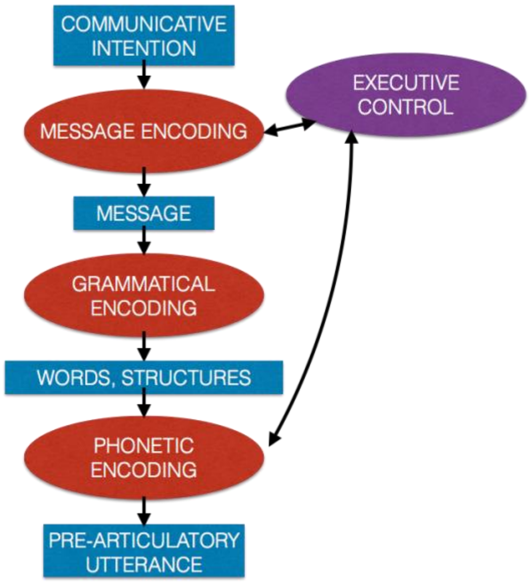

Figure 1. Model of feedforward language production.

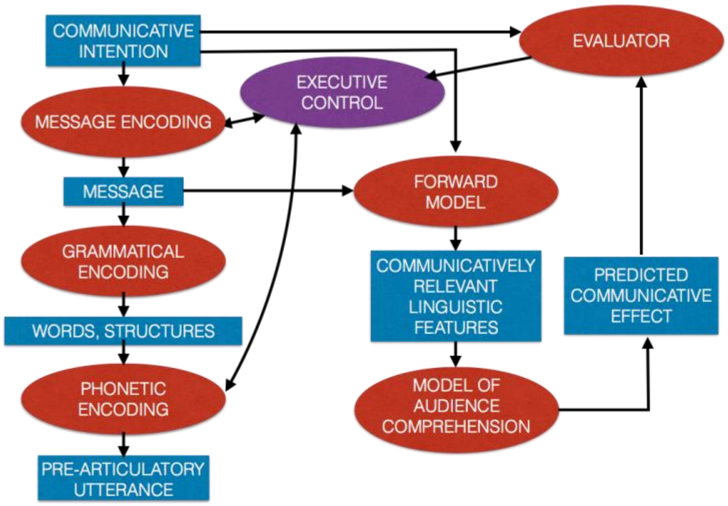

Figure 2. Architecture for audience design using forward modeling to predict and adjust based on communicative effects.

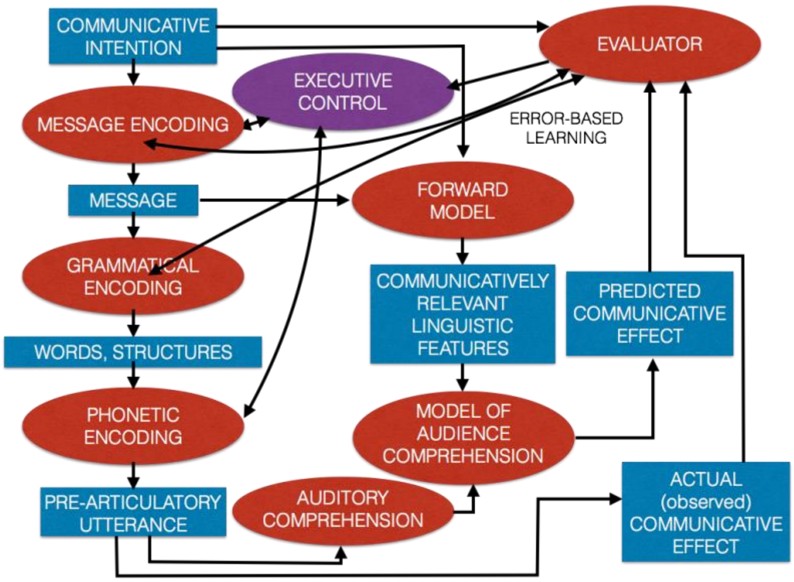

Figure 3. Complete architecture for audience design, including feedforward production, forward-model based prediction of communicative effects, and learning of new audience-design strategies.