Speech Errors as a Window on Language and Thought: A Cognitive Science Perspective

by Giulia M.L. Bencini

1. INTRODUCTION

We are so used to speaking in our native language that, under normal circumstances, we take this ability for granted. We don't assume that "talking" involves mental operations – "thinking" is where we assume that real cognitive activity resides. Under normal circumstances, native speakers only appreciate the difficulty of executing speech when something goes wrong, such as when they are temporarily unable to find the right word (tip-of-the-tongue state, TOT), or when they utter something different from what they intended to say, i.e. they produce a slip of the tongue. To the lay person, slips of the tongue are likely to be sources of amusement or embarrassment, but to the cognitive scientist they are natural experiments that reveal properties of the system that builds utterances and translates our thoughts into language.

This paper reviews the contributions of speech error data to cognitive theories of language production, because errors form part of the empirical data set upon which psycholinguistic theories of production were historically built on. With some notable exceptions (e.g. Fromkin 1971) linguists have not taken speech error data to be informative about speakers' linguistic knowledge or mental grammars. The paper wishes to rectify this and place speech errors (and more generally, language production research) back onto the linguistic data map, within a more general commitment to aim for unified theories of language representation and processing (e.g. Bencini 2013; Lewis and Phillips 2015; Branigan and Pickering 2016).

What is a speech error? A speech error (also known as slip of the tongue) is operationally defined as a deviation from a speaker's intended utterance. This definition requires knowing what a speaker's intended utterance is, which limits the classification of errors to instances in which the intended utterance is uniquely inferable from context, or when it can be verified through other means, such as in experimental paradigms that are specifically designed to elicit errors where the target utterance is either provided or constrained to an identifiable set of alternatives. Errors occur in all languages (including sign languages) and across input and output modality (e.g. reading comprehension, written production), but I will not address these other modalities here.

The focus of the paper is on spoken errors made by monolingual English native speakers, the systematic properties of those errors, and an outline of a cognitive architecture (or language production model) that explains how those errors arise during real-time language processing. Following Ferreira and Svlec (2007) I refer to the model as the consensus model for language production. Later in the paper I will review more recent challenges to the consensus model and, in light of additional experimental data, I will present a revised model for language production that aims to unify data from representation and processing. The paper aims to show that: 1) native speaker errors are rich sources of data on the nature of language representation and language use. They are, therefore, primary data for linguistics, 2) speech errors are best understood within language processing models that assemble utterances from their component units and have an overall frame-based (constructionist) architecture where frames are filled with the appropriate linguistic units: words at the syntactic level and phonemes at the phonological level.

The paper is organized as follows: In section 2, I provide an overview of the cognitive science perspective on errors. In section 3, I review processing models of language production. In order to understand the properties of speech errors and how they occur, section 4 discusses speaker error types in relation to "when and where" they occur in the processing model. Section 5 reviews some outstanding issues and current debates in language production and section 6 proposes a revised language production model to account for some of these debates.

2. A COGNITIVE SCIENCE PERSPECTIVE ON SPEAKING AND MISSPEAKING

The commonsense view of speaking alluded to earlier rests on an implicit assumption that Bock (1996) refers to as the "mind-in-the mouth assumption". This assumption is also associated with the tradition of behaviorist psychology and linguistics. "Mind-in the-mouth" refers to an output-oriented view of cognition and the idea that there is a "shallow" relationship between pre-linguistic thought and speech: no intermediate steps are required from, for example, seeing an object to naming it.

This view contrasts with the more recent perspective developed in the cognitive sciences for which generating everyday utterances is a prodigious act of linguistic creativity in the face of the computational complexity of the task. We will see, on the basis of error data and data from psycholinguistic experiments, that speaking involves selecting, retrieving and building novel combinations of units at different levels of representation, from semantics to syntax, to morphology and phonology. At each level there are different rules or regularities that speakers unconsciously follow, from rules about word order to rules about word building to rules about the sequencing of sound units. Viewed as information processing (Bock 1982), speaking is a feat of computational efficiency. The mental lexicon for the average educated adult contains anywhere between 50,000 and 100,000 words and utterance generation happens fast: during normal fluent speech, speakers can produce up to three words per second (Levelt 1999). Investigators have collected speech error corpora through direct observation, either by listening to speech and systematically writing down all the errors detected (e.g. Meringer and Meyer 1895), or by setting aside "collection times" during the day, or, more systematically, by recording speech for later examination (Garnham et al. 1981). Using this technique, Garnham et al. (1981) were able to estimate that errors in normal native adult English speakers occur at a rate of one or two errors every thousand words.

2.1 Language production data as data for linguistics

A complete cognitive theory of human language requires an account of the representations and the processes that explain how it is that humans represent and use (comprehend and produce) language. In practice, linguists in the generative tradition have proposed candidate representations (in particular with respect to syntactic and semantic structure) largely relying on explicit judgments about the grammatical (or semantic) acceptability of individual sentences – henceforth, acceptability judgments.

Over the last 25-30 years, alternative models of linguistic knowledge have developed within a class of theories that can be collectively referred to as usage-based linguistic approaches (following Langacker 1988), including Cognitive Linguistic Approaches and the family of Construction Grammars (for recent overviews, see the edited volumes by Dabrowska and Divjak (2015) for the former; Hoffmann and Trousdale (2013) for the latter). Although usage-based approaches differ in many important details, they share two important philosophical and methodological characteristics: A "cognitive commitment" to converge with data from experimental psychology, cognitive science and neuroscience and a broadened empirical basis, where (at least in principle) psycholinguistic data are to be placed on equal footing with traditional linguistic data. The revised model in section 6 is an example of a model that integrates theoretical and experimental developments within a cognitive constructionist approach to argument structure (Goldberg 1995; 2006) and a modified psycholinguistic language production model. The model is also compatible with otherwise very different linguistic neo-constructionist approaches (e.g. Borer 2005; Levin and Rappaport Hovav 2005, i.a.)

Psycholinguists over the last 50 years have examined the properties and distributions of naturalistically occurring errors (in error corpora) and set up experimental procedures to make errors more likely (e.g. Baars, Motley and MacKay 1975; Baars 1992; Warldlow Lane and Ferreira 2010). In one such technique, called SLIP (Spoonerisms of Laboratory-Induced Predisposition), speakers are presented with pairs of words on a computer screen and are asked to read them silently. Every few trials, on a tone, speakers are cued to produce the most recent word pair as quickly as possible, e.g. “barn door”. By skillfully setting up the preceding word pair sequence, (e.g. “dog bone”, “dust ball” “dead bug” “deer back” “doll bed”) the phrase “barn door” on occasion will be produced erroneously as “darn bore”. Nature breaks at its joints: the types and distributions of errors reveal how the system is built to begin with. In the error where the speaker uttered “darn bore” instead of “barn door”, there was an exchange between two word initial sounds: /b/ and /d/. The experimental set-up induced higher rates of speech errors than can be typically observed in natural speech corpora, but importantly, the mechanism is “normal” speech production (Bock and Levelt 1994). The error reveals a number of interesting properties about the system that builds words and utterances, among which that sounds like /b/ and /d/ have unit status in the minds of speakers, and that the exchange is also governed by the position that the sounds hold within larger units, e.g. syllables and words. In other words, the error is principled: we wouldn’t expect (and do not observe) an error like “noord bar” where final /n/ switches with initial /b/, or where a vowel switches with a consonant. This type of experiment, then, provides data to cognitive science, just like setting up experiments where we get atoms to collide provides data to physics (Dell 1997). Speech errors, therefore, are “primary” data in cognitive science.

Error data deserve to be considered primary sources of data in linguistics, too (following Fromkin, 1971, 1973a 1973b, and some earlier descriptive linguistic approaches to errors: e.g. Hockett 1967; Meringer and Meyer 1895). We will see that errors involve linguistic units such as phonemes, syllables, morphemes and phrases, which may be exchanged, moved around or stranded during spoken production. Errors provide evidence that the units of phonology, morphology and syntax are both representational (i.e., linguistic), and processing (i.e., psycholinguistic) units. The fact that errors are neither deliberate, strategic, nor metalinguistic makes errors desirable sources of data to consider, in the view that one source of data is not marked a priori to be “better” than others for theory development (and data may change as the field advances scientifically or technologically). If the same linguistic units are converged upon via multiple methods (e.g. native speaker judgments, spoken error corpora, language processing experiments) those units have stronger empirical support. A theory of language that can deal both with language representation and with language use, all else being equal, is a stronger theory.

3. A PROCESSING MODEL FOR LANGUAGE PRODUCTION

Cognitive scientists view the generation of an utterance as a staged process, that starts with a pre-linguistic thought, and then involves a number of translation and transduction operations within a hierarchy of levels of representation that correspond broadly to semantics, pragmatics, syntax, morphology, phonology, prosody, phonetics. The output of these linguistic levels is a speech plan that feeds into peripheral processes such as the motor commands to the articulators. Since the seminal work of Garrett (1975) researchers have worked within an overall cognitive architecture for language production which, following Ferreira and Slevc (2007) I refer to as the "consensus model".

'Consensus' does not mean unanimous agreement on the model, and I will address some of the past and presently debated issues in section 5. Unlike other areas of language research, however, there is greater agreement among researchers both in terms of what a theory of language production has to explain as well as the overall architecture of the system (Bock and Levelt 1994; Ferreira and Svlec 2007). To show how the language production model works, consider a speaker who, presented with the picture below, utters the sentence "The dogs are chasing a cat" (Figure 1). In section 4 we will examine what kinds of speech errors occur and where they occur in the model, i.e. at what stage of production.

The consensus model accounts for the planning and execution of single utterances where an utterance consists of one or more words, spoken together under a single intonational contour, or expressing a single idea.

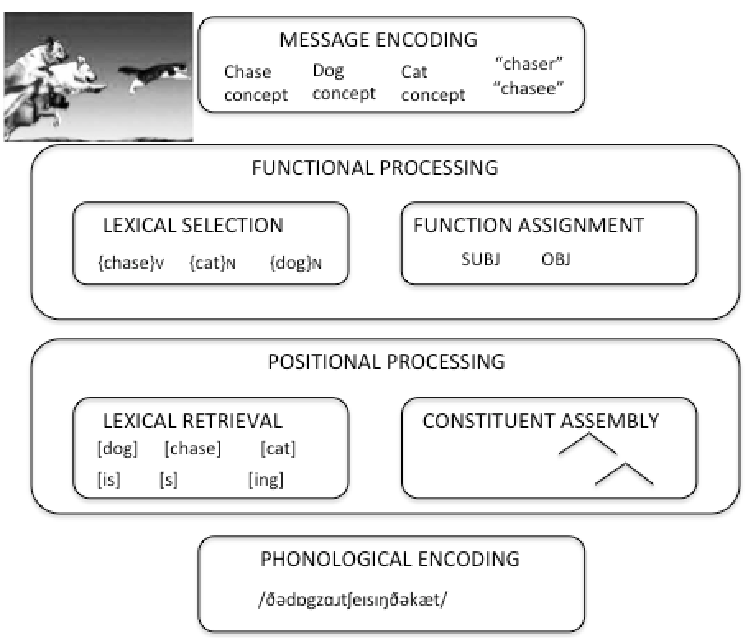

Figure 1. Schematic representation of a "consensus" model for language production. The model architecture is neutral as to whether feedback and interaction is allowed between units at different levels.

The intention to convey a meaning is called the message. The message represents conceptual and semantic information, but also information structure (e.g. *given* and *new* information), and includes aspects of *audience design*, which deals with how speakers plan their utterances with the intention of being understood by particular recipients. The next level is grammatical encoding, the powerhouse of language production; its operations are mostly invisible to a speaker's conscious awareness. Naturalistic speech error data collected by early investigators such as Fromkin (1973a; 1973b) and Garrett (1975) provided the initial evidence for a subdivision within grammatical encoding between *functional* and *positional* processes (Garrett 1975; 1980). This division was also later supported by the results of experiments with more controlled tasks that allowed researchers to tease apart the contribution of functional (semantic and syntactic) variables from the contribution of positional (phonological) variables (e.g. Bock and Warren 1985; Bock 1986a, see Bencini 2013 for reviews).

Within both functional and positional processing authors have identified two types of mechanisms: those that deal with content (lexical processes) and those that deal with structure building and ordering operations (structural processes). An important lexical process at the functional level is the selection of an abstract lexical entry called the lemma which is indexed with grammatical category information and is suitable for conveying lexical meanings to express the content of the message (e.g. {cat} from the conceptual representation of [+animate, +animal, +mammal, +domestic, +pet]). The lemma level is widely (but not unanimously, see section 5), assumed to be a modality neutral representation – that is a representation that is shared between the two language processing systems: language comprehension and language production. Lemmas contain semantic and syntactic information, but crucially no phonological information.

How are predicates represented? In the consensus model the lemma {chase} specifies that it is a verb and that it is transitive and this lexical information guides the structure building processes that assign lemmas to grammatical roles. In the utterance "The dogs are chasing the cat" the functional level selects content lemmas {dog}, {cat}, and {chase}. The {chase} lemma assigns grammatical *object* and semantic *theme* to the lemma {cat}. Representations at the functional level are left underspecified for inflectional morphology and phonology; these are retrieved at the next stage, positional processing.

Positional processing is again divided into lexical and structural processes. Lexical processing at the positional level retrieves (abstract) phonological codes called lexemes containing an overall specification of a word's syllable structure and segmental information (CV structure). Structural processes at the positional level create linearized sentence frames with slots for closed class items such as inflections, determiners and other closed class grammatical elements (tense, auxiliaries). In the example in Figure 1 there is an abstract morpheme {plural} which then undergoes phonological encoding yielding /dɒgz/). Error data (discussed in section 4) provide evidence for this ordering.

Due to the influence of linguistic theory, whereby assumptions about linguistic representations in psycholinguistics are often directly imported from theory (Bencini and Goldberg 2000; Branigan and Pickering 2016), the consensus model assumes that language production is lexically driven, in that it states that the lexical verb specifies the number and type of arguments required for sentence production. We will revisit this in sections 5 and 6.

4. Speech errors within a model of language production

When we examine errors in spoken language corpora, a remarkable property of errors stands out: errors involve phrases, words, morphemes, syllables and parts of syllables, phonemes and phonological features. Errors therefore, far from being random, reflect linguistically meaningful units. We will also see that errors occur at different stages, or levels, of the language production system outlined in 3.

This section provides examples of speech errors of different types for English, taken from classic speech error corpora. The utterances intended by the speaker (Target utterances) are given on the left of the arrow (=>) and the erroneous utterances are on the right. To illustrate where errors may occur in the processing model illustrated in Figure 1, I provide example (hypothetical) errors for the target utterance: "The dogs are chasing the cat". All other errors are from natural speech error corpora (references are given in parenthesis).

4.1. Grammatical encoding errors

4.1.1. Functional processing Errors

There are two kinds of functional processing errors: *lemma selection errors* (examples 1a – f) and *function assignment errors* (examples 2a – c).

Lemma selection errors:
(1)
a. The dogs are chasing the cat => The dogs are chasing the mouse
b. The dogs are chasing the cat => The dogs are chasing the rat
c. At low speeds it's too heavy => At low speeds it's too light (Garrett 1975)
d. Close it so it doesn't go stale => Close it so it doesn't go fresh (Fromkin 1973a)
e. Spank/paddle => spaddle (Fromkin 1973a)
f. Stomach/tummy => stummy (Fromkin 1973a)

In the examples 1a and 1b above, semantically related lemmas {mouse} and {rat} are selected instead of {cat}; these are whole word substitution errors. In 1e-f errors result in a non-word blending of two words (e.g. spladdle – a blend between spade and paddle). These non-word errors are important because they suggest, among other things, that words are assembled from combinations of smaller units (syllables, clusters, individual phonemes), during the process of language production. Phonological errors, discussed below, make this point even more evident.
Errors of lemma selection shed light onto the fact that even when speakers are producing words in isolation (e.g. naming pictures) other words in the lexicon are also active to some degree. Two dimensions that have been found to influence selection processes and to influence the probability of an error are semantics and phonology: 1) Lemma selection errors often involve words that are either semantically or phonologically related to the target; 2) Errors that are both semantically and phonologically related to the target have been found to occur at especially high rates (Dell and Reich 1981; Dell 1986).

Function assignment errors:
(2)

a. The dogs are chasing the cat => The cat is chasing the dogs
b. This seat has a spring in it => This spring has a seat in it (Garrett 1980a)
c. They must be too tight for you => You must be too tight for them (Stemberger 1982).
d. Seymor sliced the salami with a knife => Seymor sliced the knife with a salami (Fromkin 1973a)
e. I got into a discussion with this guy => I got into this guy with a discussion (Garrett 1980a)
f. Fill the car up with gas => fill the gas up with car (Dell and Reich 1981)

Function assignment errors result in so-called reverse role errors in English. They are reported to occur frequently in the speech of English speaking adults with aphasia (Thomson and Choy 2013) and in young monolingual learners. Error 2c shows that the pronouns are grammatically appropriate for the role they play in the sentence, indicating that the error occurred prior to the level in which positional features are specified, i.e. the morphological and phonological shape of the pronouns. Reverse role errors can be noticed at higher rates in experimental settings, such as structural priming paradigms (described in more detail further on), where participants previously hear or produce sentence structures of a certain type (e.g. passives) prior to describing pictures that can be described in more than one way (e.g. active: the dogs are chasing the cat, or passive: the cat is being chased by the dogs). Reverse role errors produced in priming paradigms are informative of the nature of the mapping from messages to grammatical encoding because we know exactly what picture participants are describing when they produce these errors. This responds to a general requirement in language production research that in order to tap into language production processes proper, we need to experimentally control for pre-linguistic (message level) factors (technically, we need to ensure message equivalence).

I re-examined the raw data from a syntactic priming study with English-speaking 3-year-olds reported in Bencini and Valian (2008). The study elicited active and passive sentence descriptions (reverse role errors were observed in the study but were not the focus of that study). Children produced relatively high proportions of syntactically well-formed but semantically anomalous (role-reversed) full passives for unambiguously non-reversible pictures (e.g. the knife is cut by the lemon, to describe a picture of a knife cutting a lemon; the crayon is colored by the book, for a picture of a crayon coloring a book). It is highly unlikely that, when presented with a picture of a crayon coloring a book, or a hammer cracking an egg, children interpret the picture wrong, i.e., as representing the event with reverse roles (the book is coloring a crayon; the egg is cracking a hammer). Moreover, the same children who produced reverse role errors on some trials produced semantically correct passives on other trials. The children also understood passives, as evidenced by their scores on a language comprehension test with reversible passives (e.g. show me the picture where: The frog is licked by the snake). In terms of the language production architecture in Figure 1, the error occurs in the incorrect linking of roles from the functional to the positional level. The fact that reverse role errors occur both in monolingual children and adults (albeit at different rates) shows continuity in terms representations and processing architecture across development (Bencini and Valian 2008). In sections 6, I present a revised model that accounts for reverse role errors by placing these errors neither in the message nor in the syntax, but in the linking of semantic roles in the message representation to syntactic positions in the output utterance.

4.1.2 Positional processing Errors

In the consensus model, representations at the positional level spell out the morphological (grammatical affixes and function words) and phonological information that is instead underspecified at the function level. The process of giving phonological substance to words (building lexemes) is known as phonological encoding. Positional processing errors comprise morphological errors and phonological errors.

Positional processing errors involve what we typically refer to as words, in that they include phonological information. One view of word production is that words are memorized and retrieved as wholes in comprehension and production. Another view is that words are also assembled "on the fly", as combinations of sub-lexical units. Error data in which errors see word parts come apart, move around and recombine, are consistent with the second view.

Morphological errors:
Examples 3a-d are morpheme stranding errors, whereas 3f-g are morpheme shift errors.
(3)

a. The dogs are chasing the cat => The cats are chasing the dog
b. Fancy getting your nose remodeled => Fancy getting your model renosed (Garrett 1980a)
c. She's already packed two trunks => She's already trunked two packs (Garrett 1975)
d. I hate hitching on a rainy day => I hate raining on a hitchy day (Shattuck-Hufnagel 1979)
e. I'd know one if I heard it => I'd hear one if I knew it (Garrett 1980b)
f. He goes back to => He go backs to (Garrett 1975)
g. It probably gets out a little => It probably get outs a little (Garrett 1980a)
h. Pointed out => point outed (Garrett 1980a)

Morpheme stranding errors demonstrate that stem morphemes and inflectional morphemes are dealt with separately at least during some stages of language production, and thus provide support for the division between functional and positional processing. Stranding errors 3a-c suggest that grammatical specifications such as {singular/plural} on nouns and {tense} on verbs, are part of a sentence frame and the errors are the result of the misplacement of the stem lemmas into the frame. Stranding errors in English typically involve the plural marker and past tense morphology, including irregulars, as in 3e. The stranding error in 3e shows that the abstract stem lemma {know} has switched to a position in the sentence frame which is abstractly specified for {past}, and this specification guides the insertion of the correct irregular form *knew*. The errors in 3a, 3f and 3g show that stranding occurs on abstract inflectional morphemes (e.g. {singular} and {third person singular present}) before phonological processes apply to produce the correct allomorph for the new phonological context: so we have /kaets/, not kaedz/ in 3a, and /baeks/, not /baegz/ in 3f).

Phonological errors:
Phonological encoding errors consist of anticipations, perseverations, exchanges, omissions and deletions of individual segments (phonemes), or multi-segmental material.
(4)

a. Role of simplicity => sole of simplicity (phoneme anticipation; Fromkin 1973a)
b. Fat and placid => flat and placid (anticipatory addition of a phoneme; Cutler 1988)
c. Fillmore's case grammar => Fillmore's face grammar (phoneme perseveration; Fromkin 1973a)
d. Beef noodle => beef needle (phoneme perserveration; Fromkin 1973a)
e. Left hemisphere => heft lemisphere (phoneme exchange; Fromkin 1973a)

f. Acquired dyslexias => acquired dyslexies (phoneme omission; Cutler 1988)
g. Optimal number => moptimal number (phoneme addition; Fromkin 1973a)
h. Start smoking => smart stoking (consonant cluster exchange; Fromkin 1973a)
i. Heap of junk => hunk of jeep (rhyme exchange; Dell 1981)

Phonological errors in naturalistic corpora reveal distributional patterns and strong constraints on errors. In English natural speech error corpora, two thirds of phonological encoding errors involve single segments, and approximately half of all phonological errors are either anticipations or perseverations, followed by exchanges (Stemberger 1982; Garrett 1975; Shattuck-Hufnagel 1983). When the error involves more than a single segment, it typically involves a consonant cluster, or the rhyme portion of a syllable. Errors tend to involve segments that occupy similar positions within words and within syllable structures (e.g. onsets switch with onsets, codas with codas, and nuclei with other nuclei). Phonological errors show sensitivity to word structure organization (morphological, phonological), in ways that are reminiscent of the sensitivity that lemma level errors exhibit with respect to sentence structure.

5. Debates in language production: challenges to the consensus model

Over the years, the consensus model has been challenged both at the level of lexical and structural processing. Here I will review two of those challenges. Additional discussions of past and current challenges can be found in Bock and Levelt (1994), Ferreira and Svlec (2007) and Bencini (2013).

5.1 Is language production lexically or structurally guided?

Questions about the existence of abstract, structural representations and processes operating independently of specific content have dominated much research in all areas of cognition, including language comprehension, production and acquisition. In psycholinguistics the terms "structural" and "abstract" are to be understood in contrast to lexically specific representations, and more generally the associated view that the lexical requirements of specific words (typically, verbs) are the driving force in sentence processing.

The lexicalist (also known as projectionist) view of the relationship between verb, sentence structure, and sentence meaning is one that starting with Goldberg's (1995) work in linguistics, has been challenged both on theoretical and experimental grounds (see parallel developments within generativist (neo)-constructionists approaches to argument structure: Borer 2005; Levin and Rappaport Hovav 2005 among others). Here I review experimental evidence from one language comprehension and two language production experiments with native speaking English adults. Bencini and Goldberg (2000) provided the first experimental evidence for the contribution of overall sentence structure to sentence meaning above and beyond the contribution of the main verb (in Construction Grammar terminology, they examined "constructions", defined as pairings of meaning and form). I will use the more general term "structure" to refer to the form aspects of a sentence, i.e. its (surface) syntax, and "construction" to refer to the pairing of sentence meaning and sentence structure. This is more than a terminological distinction, because Construction Grammar defines constructions as inseparable form-meaning pairings. Language production data (including reverse role errors discussed earlier, and additional data from structural priming experiments discussed below) demonstrate that form and function can dissociate at some point during the processes of producing sentences, so that production is not only meaning driven and structure is cognitively isolable (Bock and Kroch 1989). Bencini and Goldberg used a categorization task in which they pitted main verbs against sentence structures ("constructions") to determine how monolingual native English speakers would classify sentences when asked to attend to overall meaning. The stimuli were obtained by crossing four verbs (two semantically light verbs: *get*, *take* and two semantically richer verbs: *throw*, *slice*) with four different constructions (Transitive – Verb Object, e.g. "Michelle got the book"; Ditransitive – V Object1 Object2, e.g. "Chris threw Linda the pencil"; Resultative – Verb Object Result, e.g. "Nancy sliced the tire open"; Caused Motion – Verb Object Location, e.g. "Kim took the rose into the house"). Participants were asked to sort these sentences using "overall sentence meaning". They were told that there was no right or wrong answer and that the purpose of the study was to understand how people sort sentences according to meaning. Participants could thus equally sort by verb (placing all four instances of the same verb in one pile, e.g. all instances of *throw*) or by construction (placing all four instances of the same construction in one pile, e.g. all instances of the ditransitive). Categorization research outside the linguistic domain has shown that this is a stringent test to determine the role of constructions in computing sentence meaning. There is, in fact, a well-documented domain general tendency towards "one-dimensional sorting" even with categories that are designed to induce multi-dimensional sorting based on family resemblance structure (e.g. Medin, Wattenmaker and Hampson 1987). The stimuli, therefore, offered an "easy" opportunity for one-dimensional sorting based on the verb. Despite the one-dimensional sorting bias favoring verbs over constructions, results showed that speakers categorized sentences based on overall meaning by taking into account the overall argument structures of sentences, in addition to verbs. Bencini and Goldberg took these results to indicate a contribution of sentence structure to sentence meaning, independent of the contribution of the meaning of the verb.

Returning to the outline of the consensus model for language production, these results shed light onto the message level of representation: The proposal (first outlined in Bencini 2002, see also Bencini 2013; Casenhirer and Bencini 2015), is that the relational meanings that correspond to the semantics of the construction (e.g. Transitive: "X acts on Y"; Ditransitive "X causes Y to receive Z" are represented in the message independent of (and along with) more specific lexical semantic representations for verbs (e.g. the verb *chase* is associated with a lexical semantic representation that specifies the "chaser" and the one being chased, or "chasee").

Evidence for structural guidance in production (or, in constructional terminology, verb independent argument structure constructions) also comes from structural priming. Structural priming refers to the tendency of speakers to produce previously experienced sentence structures.

Priming is a powerful and sensitive method for tapping into cognitive representations and can be used to investigate linguistic representations that converge with processing (Branigan 1995; Bencini 2002; Branigan and Pickering 2016) because the priming logic allows us to draw strong inferences about the dimensions to which the cognitive architecture is sensitive. If processing of a prime stimulus influences the processing of a subsequent stimulus (the target), we can infer that the cognitive system is sensitive to the overlapping dimensions between the prime and the target. By orthogonally varying the dimensions of overlap between priming sentence and target sentence (e.g. Bock 1986b; Bock 1990; Bock and Loebell 1990; Pickering and Branigan 1998) priming has allowed researchers to study the linguistic representations at work in production, comprehension and acquisition in different populations (see Branigan and Pickering 2016, for a recent review of the main priming findings in psycholinguistics, including comprehension, production, first and second language acquisition). The original structural priming studies by Bock and colleagues (e.g. Bock 1986b; Bock and Loebell 1990) demonstrated that speakers repeated sentence structures from priming sentence to target sentence even when the prime and the target did not overlap in content words, and specifically verbs. In the classic demonstration by Bock (1986b), structural priming was demonstrated with active/passive and double object/prepositional dative constructions. Speakers were more likely to describe two-participant transitive events (e.g. a picture of dog chasing a man) with a passive if they previously heard and repeated an unrelated passive sentence with different nouns and verbs (e.g. The 747 was alerted by the airport control tower). In the original Bock (1986b) experiment, lexical overlap between primes and targets was not systematically manipulated. In a later experiment, Pickering and Branigan (1998) examined this more systematically and manipulated the amount of lexical and morphological overlap between the verbs in the prime and the target. They found an enhanced priming effect when the verb was repeated between the prime and the target, which is now referred to as the lexical boost effect in structural priming. They also found that priming was unaffected by whether tense, aspect, or number of the verb stayed the same or differed between prime and target. Verb independent priming provides strong evidence for abstract sentence representations in language production.

There are two additional experimental sets of results that provide evidence for structural guidance in language production. Konopka and Bock (2008) also used a priming paradigm to examine priming between active transitive sentences containing idiomatic (i.e., semantically non-compositional) phrasal verbs (e.g. *pull off a robbery*/*pull a robbery off*) and non-idiomatic phrasal verbs (e.g. *flip over a pancake*/*flip a pancake over*). The question was whether particle placement could be primed in idioms and non-idioms alike. They also examined whether priming would be influenced by the degree of structural flexibility of the phrasal verb (e.g. frozen idioms like *Mary's grandpa finally gave up the ghost/*gave the ghost up*, vs. flexible idioms like *The hikers broke their new boots in/ broke in their new boots*; frozen non idioms like *The new material gave off a weird smell/ *gave a weird smell off* vs. flexible non idioms like *Judy snapped on her earrings/snapped her earrings on*). On a lexicalist model, sentences with idiomatic particle verb should be stored and retrieved as lexical entries (more akin to lexical access), and not produced via the regular route assumed for sentence generation (function assignment/constituent assembly). On a structural model, in contrast, the structure building mechanisms should operate for idiomatic and non-idiomatic phrasal verb constructions. On a lexical view, therefore, idiomatic primes are predicted to exhibit reduced priming effectiveness compared to non-idiomatic primes. On a structural view, priming is predicted to be qualitatively and quantitatively similar for idioms and non-idioms. Consistent with a structural view, the results showed identical priming patterns for idioms and non-idioms (both as primes and as targets). For both idioms and non-idioms alike, structural inflexibility reduced the effectiveness of priming. This finding supports a view of language production in which the operations responsible for generating surface structure are lexically independent even for multi-word expressions that are semantically non-compositional.

A final set of results bearing on the issue of lexically versus structurally guided sentence generation comes from an error-elicitation paradigm in which speakers are asked to rapidly repeat short sentences or phrases under conditions that make errors likely. Wardlow Lane and Ferreira (2010) elicited so-called "stem exchange" errors in which words of different syntactic categories exchange positions, often stranding an inflection. For example, a speaker who erroneously says "I roasted a cook" has exchanged an intended verb *cook* with a noun *roast*, producing what appears to be an ill-formed utterance with a noun erroneously produced where a verb was intended, and vice versa. The authors selected noun and verb forms that differ in stress placement, or stress-shifting stems (the noun form REcord, vs. the verb form reCORD). If speakers can be induced to make stem exchange errors with stress-shifting stems, it is possible to determine the syntactic category of the stem. The authors elicited errors under two experimental conditions: a structurally constraining condition, which forced category membership, and a structurally neutral condition that did not. Participants heard pairs of words (e.g. REcord, hate) and were instructed to produce the words in either one of two sentence frames, prompted on the screen. In the structurally constraining transitive frame the prompt was: Use: _X_ the _Y_. The structurally neutral condition prompted the use of a conjunct clause: Use: _X_ and _Y_). The question was whether speakers would be more or less likely to produce errors as a function of the production template provided. On a structural view, speakers should be constrained by the transitive frame to produce a verb form in the first slot, and a noun form in the second slot. More errors should occur resulting in a stress-shifted stem with stress on the second syllable (i.e., resulting in a verb stem), in the transitive frame than in the conjunct frame. On a lexical view, the rate of such errors should not differ across frames. Results confirmed the predictions of the structural view, suggesting that even in the generation of speech errors, constraints from lexically independent sentence level representations operate in sentence production.

5.2 Are representations modality general or modality specific? Is the lemma/lexeme distinction necessary?

In the consensus model, the lemma is a modality general, abstract representation shared between comprehension and production, whereas the lexeme is a modality specific representation containing phonological information. Starting with an initial challenge in Caramazza (1997), Caramazza and colleagues have proposed a model that denies the existence of an abstract lemma in favor of a model that only recognizes modality specific representations (one for each modality) that combine syntactic, semantic and phonological information. With respect to language production, the modality specific representation specifies all of a word's properties (syntactic, semantic and phonological) and these operate in parallel to determine a word's position in a linearized string.

There is strong neuropsychological evidence for the modality specific nature of representations in the form of "double dissociations" in individuals with aphasia – the most striking being the double dissociations within a single individual. For example, Caramazza and colleagues examined a patient with aphasia who showed a noun/verb dissociation whereby he produced more errors in a naming task for nouns than for verbs in the spoken modality, and the opposite pattern in the written modality (Rapp and Caramazza 2002). The debate over whether the language processing systems operate on modality specific representations or whether there are modality neutral representations that are shared between comprehension and production has consequences for the relationship between linguistic theories and cognitive science. A strong version of the cognitively and neurologically separate representations view undermines the relevance of linguistic representations gleaned from more traditional linguistic means (e.g. native speaker judgments, corpus data) to their involvement in understanding and producing language. This debate is still open in neuropsychology. It poses a concern to all theories of language that aim to be theories of the mental representation of language (i.e. all cognitive theories – whether mainstream generativist, functionalist or usage-based). One line of research that can address these questions empirically is to conduct studies that compare results from tasks within one modality (e.g. production) with results across modalities (comprehension to production).

One such study was conducted by Bock and colleagues (2007) who systematically compared comprehension-to-production-priming with production-to-production-priming. Participants described pictures after listening to sentences in different structures (active/passive; ditransitive/prepositional dative), which, crucially, they did not repeat. Because priming involves the reusing of previous representations/processes, if representations are modality specific, this predicts that comprehension-to-production priming should not mirror production-to-production priming. If representations are abstract and shared between comprehension and production, similar amounts of priming are expected across modalities. Results showed remarkably similar cross-modal priming, both in terms of magnitude and temporal duration, to previous within-modality priming results. More cross-modality priming studies need to be conducted in different populations, but I take the existence of cross-modal syntactic priming from comprehension to production to be strong evidence for modality general representations that are shared between comprehension and production. This is a very important result because it provides empirical support for a greater convergence between linguistic theories and theories of language processing.

6. A REVISED MODEL FOR LANGUAGE PRODUCTION

The body of experimental evidence reviewed in section 5 which includes error data, converging data from language comprehension and language production and theoretical insights from linguistics, allows us to propose a modified language production model that accounts for verb-independent priming effects. Returning to the example utterance in Figure 1 ("The dogs are chasing the cat") we saw that the consensus model requires the selection of the lemma for the verb *chase*, which, in turn, makes available (projects) its argument structure preferences represented within the lemma level. The essence of the revised model is that it allows for direct mappings from messages to grammatical encoding independent of the verb lemma (i.e. it recognizes the contribution of argument structure constructions to sentence meaning). In the variant of the model, sentence production operates on representations that segregate lexical and structural information (for a more detailed discussion see Bencini 2013). In the example above, the transitive active sentence frame is not activated solely by the selection of the verb *chase*, rather it results from the combination of {chase} and an abstract (lexically-independent) mapping from the message to a sentence structure with an associated general semantics roughly corresponding to X acts on Y (An abstract argument structure construction in Construction Grammar). It is important to note that the revised model does not deny the importance of lexical (or item specific) processes in language use because lexical and structural processing must converge to produce well-formed utterances. The verb contributes lexically-specific aspects of meaning; but crucially, structural processes do not rest upon prior lexical access and retrieval. We see this clearly in English expressions where there is a mismatch between the number of thematic roles associated with the verb lemma and grammatical positions in the syntax as in Goldberg's classic example: "She sneezed the foam off the cappuccino" where *sneeze* appears with a direct object when it is prototypically an intransitive verb (see Goldberg 1995; 2006). Under the structurally guided view, the lexically independent construction contributes the two additional argument roles and the overall cause-to-receive interpretation, while the predicate contributes the means component of the event.

7. Summary and some concluding remarks

In this paper, I have provided an account of native speaker errors from the point of view of cognitive science, which is interested both in how language is represented and also in how it is used. I have reviewed evidence that demonstrates that errors exhibit systematic properties and remarkably reflect linguistic units employed by linguists to characterize native speaker competence. Errors are highly rule governed and reflect the architecture and the processes of the language production system that generates utterances. Errors have traditionally been dismissed as data for linguistics. They shouldn't be, especially if one's goal is to provide broader theories of language that encompass both representation and use.

Works cited

Baars B.J., M.T. Motley and D.G. MacKay, 1975, "Output Editing for Lexical Status in Artificially Elicited Slips of the Tongue", *Journal of Verbal Learning and Verbal Behavior* 14, pp. 382–391.

Baars B.J., 1992, *Experimental Slips and Human Error. Exploring the Architecture of Volition*, Plenum Press, New York.

Bencini G.M.L., 2002, *The Representation and Processing of Argument Structure Constructions*, PhD Dissertation, University of Illinois at Urbana-Champaign.

Bencini G.M.L., 2013, "Psycholinguistics". In T. Hoffmann and G. Trousdale (eds.), *Construction Grammar Handbook*, Oxford University Press, Oxford, pp. 379–396.

Bencini G.M.L. and A.E. Goldberg, 2000, "The Contribution of Argument Structure Constructions to Sentence Meaning", *Journal of Memory and Language* 43, pp. 640–51.

Bencini G.M.L. and V. Valian, 2008, "Abstract Sentence Representations in 3-year Olds: Evidence from Comprehension and Production", *Journal of Memory and Language* 59, pp. 97–113.

Bock J.K. 1982, "Toward a Cognitive Psychology of Syntax: Information Processing Contributions to Sentence Formulation", *Psychological Review* 89, pp. 1–47.

Bock J.K., 1986a, "Meaning Sound and Syntax: Lexical Priming in Sentence Production", *Journal of Experimental Psychology: Learning Memory and Cognition* 12, pp. 575–586.

Bock J.K., 1986b, "Syntactic Persistence in Language Production", *Cognitive Psychology* 18, pp. 355–387.

Bock J.K., 1996, "Language Production: Methods and Methodologies", *Psychonomic Bulletin & Review* 4, pp. 395–421.

Bock J. K. and A. S. Kroch, 1989, "The Isolability of Syntactic Processing", In G.N. Carlson and M.K. Tanenhaus (eds.), *Linguistic Structure in Language Processing*, Kluwer, Dordrecht.

Bock K. and W. Levelt, 1994, "Language Production: Grammatical Encoding", in M. A. Gernsbacher (ed.), *Handbook of Psycholinguistics*. Academic Press, San Diego, pp. 945–84.

Bock K. and R. Warren, 1985, "Conceptual Accessibility and Syntactic Structure in Sentence Formulation", *Cognition* 21, pp. 57–67.

Bock K., G. Dell, F. Chang and K. Onishi, 2007, "Persistent Structural Priming from Language Comprehension to Language Production", *Cognition* 104, pp. 437–58.

Borer, H., 2005, *In Name Only. Structuring Sense*, vol. 1. Oxford University Press, Oxford.

Branigan H., M. Pickering, S. Liversedge, A. Stewart and T. Urbach, 1995, "Syntactic Priming: Investigating the Mental Representation of Language", *Journal of Psycholinguistic Research* 24, pp. 489–506.

Branigan H. and M. Pickering, 2016, "An Experimental Approach to Linguistic Representation", *Behavioral and Brain Sciences*, November, pp. 1–73.

Bresnan J. and R. Kaplan, 1984, "Grammars as Mental Representations of Language," in W. Kintsch and P. Polson (eds.), *Methods and Tactics in Cognitive Science*, Erlbaum, Hillsdale, pp. 103–135.

Caramazza A., 1997, "How Many Levels of Processing are There in Lexical Access?", *Cognitive Neuropsychology* 14, pp. 177–208.

Casenhiser, D. and G.M.L. Bencini, 2015, "Argument Structure Constructions", In E. Dabrowska & D. Divjak (eds.), *Handbook of Cognitive Linguistics*, de Gruyter-Mouton, pp. 579–594.

Cho S. and C.K. Thompson, 2010, "What Goes Wrong During Passive Sentence Production in Agrammatic Aphasia: An Eyetracking Study", *Aphasiology* 24, pp. 1576–1592.

Cutler A., 1988, "The Perfect Speech Error", in L.M. Hymann and C. N. Li (eds.), *Language Speech and Mind: Studies in Honor of Victoria A. Fromkin*, Routledge, London, pp. 209–223.

Dell G.S., 1986, "A Spreading-Activation Theory of Retrieval in Sentence Production", *Psychological Review* 93, pp. 283–321.

Dell G.S., 1995, "Speaking and Misspeaking", in L. Gleitman and M. Liberman (eds.), *Language: An Invitation to Cognitive Science*, MIT Press, Cambridge, pp. 183–208.

Dell G.S. and P. Reich, 1981, "Stages in Sentence Production: An Analysis of Speech Error Data", *Journal of Verbal Learning and Verbal Behavior* 20, pp. 611–629.

Ferber R., 1993, *Vie valide sind versprechersammlungen?*, Peter Lang, Bern.

Fromkin V.A., 1971, "The Nonanomalous Nature of Anomalous Utterances", *Language* 47, pp. 27–52.

Fromkin V.A., 1973a, "Appendix", In V. A. Fromkin (ed.), *Speech Errors as Linguistic Evidence*, Mouton, Den Hague, pp., 243–269.

Fromkin V.A., 1973b, "The Nonanomalous Nature of Anomalous Utterances", in V. A. Fromkin (ed.) *Speech Errors as Linguistic Evidence*, Mouton, Den Hague, pp. 215–241.

Fromkin V.A., 1973c, *Speech Errors as Linguistic Evidence*, Mouton, Den Hague.

Ferreira V. and R. Slevc, 2007, "Grammatical Encoding", in G. Gaskell (ed.), *The Oxford Handbook of Psycholinguistics*, Oxford University Press, Oxford, pp. 453–69.

Garnham A., R.C. Shillcock, G.D.A. Brown, A.I.D. Mill, and A. Cutler, 1981, "Slips of the Tongue in the London-Lund Corpus of Spontaneous Conversation", *Linguistics* 19, pp. 805–817.

Garrett M., 1975, "Syntactic Processes in Sentence Production", in G. Bower (ed.), *Psychology of Learning and Motivation: Advances in Research and Theory*, Academic Press, New York, pp. 133–177.

Garrett M., 1980a, "Levels of Processing in Sentence Production" in B. Butterworth (ed.), *Language Production* Vol. 1, Academic Press, London, pp. 177–220.

Garrett M., 1980b, "The limits of accommodation", in V.A. Fromkin (ed.), *Errors in Linguistic Performance*, Academic Press, New York.

Goldberg A.E., 1995, *Constructions: A Construction Grammar Approach to Argument Structure*, University of Chicago Press, Chicago.

Goldberg, A.E, 2006, *Constructions at Work: The Nature of Generalization in Language*, Oxford University Press, Oxford.

Hockett C.F., 1973, "Where the Tongue Slips There Slip I", in V.A. Fromkin, *Speech Errors as Linguistic Evidence*, Mouton, Den Hague, pp. 97–114.

Langacker R.W., 1988, "Usage-based Model" In B. Rudzka-Ostyn (ed.), *Topics in Cognitive Linguistics*, John Benjamins, Amsterdam, pp. 127–161.

Levelt W., 1999, "A Blueprint of the Speaker", in C. Brown and P. Hagoort (eds.), *The Neurocognition of Language*, Oxford University Press, Oxford, pp. 83–122.

Levin, B. and M. Rappaport Hovav, 2005, *Argument Realization*, Cambridge University Press, Cambridge; New York.

Lewis S. and C. Phillips, 2015, "Aligning Grammatical Theories and Language Processing Models", *Journal of Psycholinguistic Research* 44, pp. 27–46.

Medin D.L., W.D. Wattenmaker and S.E. Hampson, 1987, "Family Resemblance, Conceptual Cohesiveness, and Category Construction", *Cognitive Psychology* 19 pp. 242–279.

Meringer R. and Mayer K., 1895, *Misspeaking and Misreading: A Psycholinguistic Study*, Goschense Verlagsbuchhanlung, Stuttgart.

Rapp B. and Caramazza A., 2002, "Selective Difficulties with Spoken Nouns and Written Verbs: A Single Case Study", *Journal of Neurolinguistics* 15, pp. 373–402.

Shattuck-Hufnagel S., 1979, "Speech Errors as Evidence for a Serial Order Mechanism in Sentence Production", in W. E. Cooper and E.C.T. Walker (eds.), *Sentence Processing: Psycholinguistic Studies Presented to Merrill Garrett*, Erlbaum, Hillsdale, pp. 295–342.

Shattuck-Hufnagel S., 1983, "Sublexical Units and Suprasegmental Structure in Speech Production Planning" in P.E. MacNeilage (ed.), *The Production of Speech*, Springer, New York, pp. 109–136.

Stemberger J.P., 1982, "The Nature Of Segments in the Lexicon: Evidence from Speech Errors", *Lingua* 56, pp. 235–259.

Stemberger J.P., 1983, "Inflectional Malapropisms: Form-Based Errors in English Morphology", *Linguistics* 21, pp. 573–602.

Giulia M. L. Bencini is an Associate Professor in the Department of Linguistics and Comparative Cultural Studies at Ca Foscari University, Venice. Before joining the faculty at Ca Foscari, she was an Associate Professor in Speech-Language Pathology and Audiology at Hunter College, City University of New York. Her research interests are broadly in cognitive approaches to language knowledge and use in children, adults and individuals with language impairments. This work is situated within a broader bio-psycho-social approach to human functioning which views language as a specialized mental function and communication as an activity (ICF: International Classification of Functioning, disability and health).

giulia.bencini@unive.it