Against stored abstractions:

A radical exemplar model of language acquisition

PREPRINT: NOT PEER REVIEWED

Ben Ambridge

University of Liverpool

ESRC International Centre for Language and Communicative Development (LuCiD)

Running head: EXEMPLAR MODEL

Keywords: *child language acquisition; exemplar account; instance-based; memory-based; word meaning; syntax; morphology; phonology; n-grams; construction; abstraction*

Acknowledgements: Ben Ambridge is Professor in the International Centre for Language and Communicative Development (LuCiD) at The University of Liverpool. The support of the Economic and Social Research Council [ES/L008955/1] is gratefully acknowledged. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no 681296: CLASS).

Against stored abstractions: A radical exemplar model of language acquisition

The goal of this paper is to make the case for a radical exemplar account of child language acquisition, under which unwitnessed forms are produced and comprehended by on-the-fly analogy across multiple stored exemplars, weighted by their degree of similarity to the target with regard to the task in hand. Across the domains of (1) word meanings, (2) morphologically inflected words, (3) *n*-grams, (4) sentence-level constructions and (5) phonetics and phonology, accounts based on independently-represented abstractions (whether formal rules or prototype categories) fail for two reasons. First, it is not possible to posit abstractions that delineate possible and impossible forms; e.g., that (a) rule in pool *tables* and data *tables*, but rule out chairs, (b) rule in "mist" and rule out "missid" as the phonological form of the past tense of *miss*, (c) rule in both *John feared Bill* and *John frightened Bill* but rule out **John laughed Bill*. Second, for each domain, empirical data provide evidence of exemplar storage that cannot be captured by putative abstractions: e.g. speakers prefer and/or show an advantage for (1) exemplar variation even within word-meaning "category boundaries", (2) novel inflected forms that are similar to existing exemplars, (3) *n*-grams that have occurred frequently in the input, (4) SVO sentences with *he* as SUBJECT and *it* as OBJECT and (5) repeated productions of "the same" word that are phonologically similar or, better still, identical. An exemplar account avoids an intractable lumping-or-splitting dilemma facing abstraction-based accounts and provides a unitary explanation of language acquisition across all domains; one that is consistent with models and empirical findings from the neuroimaging literature.

Against stored abstractions: A radical exemplar model of language acquisition

0. Introduction: A rare, but misplaced, consensus in language acquisition research

Most, perhaps all, mainstream theories of child language acquisition share a common assumption: Adult knowledge of language includes stored abstractions such as [VERB] [NOUN] and [SUBJECT]; and language acquisition therefore involves forming and/or mastering these abstractions. The goal of the present article is to argue that this assumption is misplaced, and to present the case for a radical alternative: Adult knowledge of language consists of nothing but stored exemplars; and language acquisition therefore involves simply storing these exemplars.

Child language acquisition is often seen as a highly polarized field, dominated by an all-encompassing nature-nurture debate (e.g., Valian, 2014). But, when it comes to the question of stored abstractions, there is widespread agreement. While they differ on the details, all sides agree that adult speakers possess stored linguistic abstractions of some kind. Indeed, though some stored abstractions are controversial (e.g., independently represented passive or *wh*-question constructions), others – such as the category [VERB], and some kind of abstract representation of canonical word order (e.g., English [SUBJECT] [VERB] [OBJECT]) – are agreed upon by virtually all theories, with disagreements revolving solely around timing: Are these abstractions present from birth (e.g., Pinker, 1989; Wexler, 1998), formed very quickly (e.g., Gertner, Fisher & Eisengart, 2006), or formed more gradually (e.g., Tomasello, 2003), as posited by *generativist-nativist*, *early-abstraction*, and *usage-based-constructivist* accounts respectively? The widespread theoretical appeal of such abstractions is obvious; they offer an explanation of the fact that speakers are able to produce and comprehend sentences that they have never heard before (as in Chomksy's famous example *Colourless green ideas sleep furiously*).

But we should not be so easily seduced. In attempting to rebut these stored-abstraction accounts in favour of a radical exemplar account[1], I develop three lines of argument. The first holds that the apparent explanatory power offered by these stored abstractions is illusory: There is simply no way to formulate a linguistic abstraction – be it the category [VERB], SVO word order, or the word *table*– that rules *in* all the currently-unwitnessed exemplars that are permissible, but rules *out* those that are not (see Ramscar & Port, 2016, for discussion of the difficulties inherent in positing any kind of discrete units in language representation). The second line of argument is that across a wide range of acquisition domains, the evidence for storage of individual exemplars is overwhelming. The third rejects the (often-implicit) claim that stored-abstraction accounts are parsimonious while exemplar accounts are implausible in terms of the required memory capacity. In developing these three arguments, I consider representations at five levels: **(1) word meanings, (2) morphologically inflected words, (3) *n*-grams (primarily at the word level), (4) sentence-level constructions and (5) phonetics and phonology**.

I end by presenting the alternative: an exemplar account under which unwitnessed forms are produced and comprehended by on-the-fly analogy across multiple stored exemplars, weighted by their degree of similarity to the target *with regard to the task in hand*. For example, consider the single exemplar *Mummy kissed her.* If the task in hand is to generate the appropriate phonological form for the past-tense form for *wish*, phonological similarity is relevant (e.g., *kiss*→*kissed*, so *wish*→*wished*). If the task in hand is to express

[1] The use of the term *exemplar account* is intended to differentiate the present account from merely *exemplar-**based*** accounts, which allow for the possibility that stored exemplars are merely a stepping stone on the path towards more abstract representations.

who did what to whom in a hugging scenario, one relevant dimension is semantic similarity, in terms of the ordering of semantic roles (e.g., [AGENT] [ACTION] [PATIENT).

Although much work remains to be done to flesh out the details, various types of computational models that are broadly consistent with this approach have enjoyed important empirical successes, while a model based around the stored abstractions posited almost universally by verbal accounts of language acquisition is literally a non-starter: Such models are not found in the literature, because they do not, and cannot, work.

First, it will be useful to consider a very simple example that illustrates the principles of an exemplar account, and contrasts it against rival stored-abstraction approaches of both the generativist-nativist and usage-based-constructivist variety. Suppose that an English-speaking 4-year-old is taught a novel verb ("Look, tamming!") to describe the motion of a toy block bouncing and spinning on a suspended rope (e.g., Brooks & Tomasello, 1999), and then describes a scene in which a doll bounces and spins on the rope by saying *She's tamming*. Stored-abstraction accounts offer the following explanations.

Generativist-nativist accounts (e.g., Wexler, 1998) assume that children are born with the syntactic categories [SUBJECT] and [VERB], and with the knowledge that these two categories can be combined to yield a phrase or sentence. Furthermore, a four-year-old will have (long ago) set the specifier-head parameter, which determines (amongst other things) whether the ordering of these categories in the target language is SV (e.g., English) or VS (e.g., Welsh, Maori). (We ignore here the additional complication introduced by the auxiliary *is*). This knowledge, together with the knowledge that *She* and *tamming* are respectively a SUBJECT and – based on its meaning and morphology –a VERB allows the speaker to generate *She's tamming*.

Usage-based constructivist accounts (e.g., Tomasello, 2003) assume that children generalize across input sentences (e.g., *She's dancing, She's playing, She laughing*) to form, first, slot-and-frame patterns (e.g., *She's [ACTION]ing*) and, later, fully-abstract sentence level constructions (e.g., [SUBJECT] [VERB]). A four-year-old producing *She's tamming* would normally be assumed to have generated the utterance by inserting the relevant lexical items into the abstract construction, though a much younger child (around 2;0, say) would normally be assumed to have used the lower-level slot-and-frame pattern.

Both of these accounts assume that speakers store free-standing linguistic abstractions. Indeed, although the acquisition processes are very different, the mature abstractions – [SUBJECT] [VERB] phrase structure and a [SUBJECT] [VERB] intransitive construction – are almost indistinguishable. An exemplar account, on the other hand, assumes no such stored abstractions. Rather, learners store concrete exemplars, each including the surface form along with its understood meaning and contextual details, and produce (and comprehend) novel utterances in real time by analogizing across these stored exemplars on the basis of similarity. The basis of this similarity and the precise form of the analogy depend on the particular implementation under consideration; and, as we will see later, for a few circumscribed acquisition problems, there exist fully-specified computational implementations of an exemplar account. For the purposes of the present toy example, however, it will suffice to say that the child generates *She's tamming* to express the message 'Discourse-old female undergoes spinning action' by analogy across utterances with very similar semantic/functional properties (e.g., 'Discourse-old female undergoes bouncing action' = *She's bouncing*; 'Discourse-old female undergoes spinning action' = *She's spinning*), more distant analogy across utterances with somewhat similar semantic/functional properties (e.g., 'Discourse-new female undergoes spinning action' = *Sue's spinning*; 'Discourse-old male undergoes bouncing action' = *He's spinning*), and the utterance in which the novel verb was trained ('Toy block undergoes bouncing+spinning action' = *Look, tamming!*), as summarized in Table 1.

***Table 1.* Production of the novel utterance *She's tamming* by analogy across stored exemplars with similar semantic/functional properties ("message") to the target**

Stored message	Stored utterance
Toy block undergoes bouncing+spinning action	*Look, tamming!*
Discourse-old male undergoes bouncing action	*He's spinning*
Discourse-new female undergoes spinning action	*Sue's spinning*
Discourse-old female undergoes bouncing action	*She's bouncing*
Discourse-old female undergoes spinning action	*She's spinning*
Target: Discourse-old female undergoes spinning action	**Output: *She's tamming***

It is important to stress from the outset that, as this toy example illustrates, an exemplar account does not posit that speakers are restricted to producing words, phrases or sentences from a fixed repertoire: At all levels, novel combinations can be generated by analogy as soon as the learner has stored, in principle, a single relevant exemplar (e.g., *go*→**goed* by phonological analogy with *show*→*showed*). Hence, an exemplar model (at least in some possible implementations) actually predicts earlier comprehension and production of novel forms than a constructivist account under which abstractions (e.g., the [SUBJECT] [VERB] construction discussed above) are acquired only when children reach a critical mass of exemplars across which to generalize (e.g., Marchman & Bates, 1994).

It is also important to stress from the outset that an exemplar approach does not entail abandonment of constituent structure. An important insight captured by traditional generative linguistics is that linguistic generalizations operate on constituents, rather than particular words. For example, whatever generalization one posits for forming a question with *Bill* must also apply for *the man*, *the man who is tall* and so on (e.g., Chomsky, 1980; Crain & Nakayama, 1987):

Bill is happy → Is Bill happy?
The man is happy → Is the man happy?
The man who's tall is happy → Is the man who's tall happy?

Generativist and Constructivist accounts alike capture constituent structure by positing the existence of stored abstractions like [SUBJECT] or [NOUN PHRASE] (e.g., *Bill; the man; the man who is tall*). Exemplar accounts do not posit stored abstractions; rather, constituent structure is inherent in the exemplars across which analogies are drawn. The same analogical processes that allow the system to analogize across semantically-similar utterances with *He, She and Sue* (Table 1) allow it to analogize across semantically-similar utterances with *Bill*, *the man* and *the man who's tall.*

1. Word meanings

Probably the simplest form of abstraction that is posited by most current theories of language acquisition is the monomorphemic word (e.g., *table*). Indeed, this abstraction is – it would seem – so simple and self-evident that it is easy to forget that it is an abstraction at all. But, of course, an abstraction it is. At the phonetic level, the idealized form *table* is an abstraction across all the different pronunciations that have been witnessed (e.g., by different speakers, by the same speaker on different occasions, and so on). Discussion of this level of abstraction is reserved for Section 5 (phonetics and phonology). Here, we focus on the semantic level, where the word *table* is an abstraction that "maps onto" or somehow "stands for" some entities in the physical world, other entities that share certain salient properties (e.g., the geological water table), photographs of those entities (or renderings in abstract art), toy versions for use in a doll's house, a talking table in a children's cartoon, and so on.

Notice that no distinction is drawn here between learning *table* the *word*, and *table* the *concept*. This is because there is no meaningful distinction to draw (Ramscar & Port, 2015). As the examples above show, the wide variety of entities that English speakers refer to as *tables* share no defining characteristic, other than being referred to as *tables* by English speakers (even if we set aside more obviously problematic cases like data *tables* or multiplication *tables*). Already, then, the cracks in the idea of an abstract meaning representation for *table* are starting to show. There is no way to define the abstraction that rules in everything that an English speaker could conceivably refer to as a *table*, and rules out everything that she could not.

Cognitive psychologists abandoned long ago (e.g., Rosch & Mervis, 1975) the idea that categories in the world are "rule-based" (Smoke, 1932) (although of course humans are *capable* of making rule-based categorizations in experiments or, for example, the legal system). Any conceivable rules for defining a *table* (e.g., "has legs"; "used for eating"; "made of wood, metal or plastic", "waist height") can be easily dismissed with counterexamples (e.g., an empty beer barrel used as a table at a bar; a pool table; an origami paper table glued to the ceiling as part of an art exhibit, and so on). And even if we could define a *table* in these terms (e.g., "has legs and a flat surface"), this only shifts the problem elsewhere (how do we define a *leg*, a *surface*, *flat*?; Ramscar & Port, 2015)

A *prototype category* meaning for *table* fares better, but is still highly problematic. On this view, speakers average across every inferred referent of *table* that they have encountered to form an abstract fuzzy, probabilistic representation of the category labelled by the word *table* (in the same way that they are argued to abstract across instances of VERBs or SVO constructions to form other linguistic abstractions). New uses of *table* are interpreted (in comprehension) or coined (in production) with reference to this abstraction; i.e., people are faster and more willing to accept novel, previously unseen items as *table*s if they are very similar to the *table* prototype. Indeed, a nonlinguistic categorization study with dot patterns (Posner & Keele, 1970) went further in demonstrating that prototypical but novel patterns were more likely to be assigned to the relevant category than were patterns that were less typical of the category (though still consistent with it), and that had been shown during training.

However, the prototype category approach suffers from three problems. The first is an empirical problem: In laboratory categorization studies, if the stimuli to be categorized are matched on prototypicality, participants show an advantage for previously-seen items (e.g., Zaki, Nosofsky, Stanton & Cohen, 2003). This is a problem for prototype models, which store only the prototype, not the individual items. Furthermore, prototype effects for unseen items, of the type demonstrated by Posner and Keele (1970) are also yielded by all but the very simplest type of exemplar model (e.g., Medin & Schaffer, 1978; Mack, Preston & Love, 2013; see Smith & Minda, 2000; Zaki et al, 2003, for reviews). This is because a prototypical item – even one that has not been seen before – is highly similar to a large number of stored exemplars. Consequently, exemplar models that analogize to new items on the basis of similarity across a number of stored exemplars (e.g., *k*-nearest-neighbour models) yield prototypicality effects for unseen items, without a stored representation of the prototype. (A maximally simple exemplar model that analogizes on the basis of the *single* nearest neighbour cannot yield this effect). In sum, as Love (2013: 348) puts it "By and large, exemplar models can mimic all the behaviors of prototype models, but the opposite is not true", including when accounting for fMRI data obtained from participants performing categorization tasks (e.g., Nosofsky, Little & James, 2012; Mack et al, 2013). It is therefore surprising that virtually all accounts of word learning (and child language acquisition more generally) have – whether implicitly or explicitly – opted for the latter.

The second is an in-principle problem: For real-word putative word categories like *table* (as opposed to dot-pattern or abstract-shape categories in a lab-based classification task), there is no way to define the prototype: Do speakers just have a single prototype meaning for *table*, that includes domestic dining tables, beer-barrel bar tables and fold-down aeroplane tables (lumping), or a separate prototype for each (splitting)? The lumping approach is unworkable because some (would-be) categories have internal structure. For example, spoons are generally small and metal or large and wooden; but nobody would define a prototypical spoon as one that is of intermediate size and made out of an intermediate wood-metal material (example from Love, 2013). The splitting approach is unworkable, because there is no principled way to stop splitting. Do we have a single prototype of a domestic dining table, or subtypes of wooden and metal tables, or of vintage and modern tables? (or, for that matter, of data tables and of multiplication tables?).

The third problem for prototype approaches is an in-practice problem. Prototype categories might be useful for dot-pattern classification tasks, but they are useless in the real world. Suppose a listener is asked if a (beer-barrel bar-room) *table* is free, or to put away her (aeroplane tray) *table*. The likelihood of communicative success is determined not by the extent to which a general prototype invoked for the listener by the word *table* overlaps with the tables present (which is probably not by much). Rather, it is determined by the listener's inference regarding the speaker's most likely meaning of *table*, *in that particular context*. Which dimensions are relevant when deciding whether or not something can be called a *table* are not fixed, but depend on the speaker's goals. If the relevant dimension is function – I need somewhere to put my glass – I am quite happy to call an upturned beer barrel a *table*. If the relevant dimensions are aesthetic, for example when looking at a picture in a gallery, I am quite happy to refer to a particular arrangement of paint on canvas as a *table*; but I would not put my drink on it.

In an important sense, the three problems are just different ways of saying the same thing: A word like *table* does not have a standalone, prototypical or central meaning that is devoid of context. For any individual language learner, a word like *table* has as many different meanings as the leaner has been in situations in which she has interpreted a speaker's meaning of *table*. That is, for any individual language learner, a word like *table* has not one meaning, but thousands – one for each *table* that she has encountered. This is what an exemplar account of word-meaning looks like.

An advantage of an exemplar account of word meaning[2] is that it sidesteps altogether a problem that has given rise to a whole sub-discipline of language acquisition research: how children disambiguate homophones (e.g., river *bank* vs money *bank*), particularly noun/verb homophones (e.g., John likes *fish*; John can *fish*; Pinker, 1987) (see e.g., Conwell, 2017, for a review). But the problem only arises because researchers have assumed (in most cases presumably implicitly) a stored-abstraction, prototype-based model of word meaning, under

[2] Fodor and Lepore (1996) argue that neither exemplar nor prototype models can account for compositionality; e.g., for the fact that "a goldfish is a poorish example of a fish, and a poorish example of a pet, but it's quite a good example of a pet fish" (p.262). In-principle arguments aside, Kim and Baldwin (2006) demonstrated that, in practice, an exemplar model (based on the Tilburg Memory Based Learner, TiMBL), achieves good performance at interpreting compounds of this type (see also

which a word like *bank, fish* or *table*, "should" have a single meaning. The problem simply does not arise assuming an exemplar model, under which every heard exemplar of the word *bank, fish* or *table* is stored with its meaning, *as understood in that individual situation* (e.g., Erk & Padó, 2010). An exemplar model also avoids drawing an unprincipled distinction between "true" homophones (e.g., river *bank* vs money *bank*) and extensions of a "prototypical sense". Is a water *table* the "same kind" of table as a bar-room beer-barrel *table* and a journal article results *table*? Who can say? The advantage of an exemplar model is that we don't have to. *Table* has not one meaning, or two meanings that somehow require disambiguation (e.g, *table* the noun and *table* the verb), but thousands.

So ingrained is the (implicit) prototype model of word meaning that few, if any, experimental studies have investigated an exemplar account. However, two observed experimental phenomena – although the study authors do not interpret them in these terms – would seem to constitute support for an exemplar model.

First, Gordon, McGregor, Waldier, Curran, Gomez and Sameulson (2016) found that, when taught novel words for novel objects, 3-5 year-old children were able to select the correct phonological form – from a choice of either three or four – when shown the novel object a year later. While this finding does not constitute direct evidence for an exemplar over prototype account, it challenges the widely-held assumption that young language learners do not have the memory capacity that would be necessary for exemplar learning.

Second, a robust finding in the word learning literature (as well as learning in other domains) is that a label is more easily generalized to unseen items if the exemplars originally used to train the label were highly variable than highly similar (e.g., Oakes Coppage, & Dingel, 1997; Loewenstein & Gentner, 2001; Childers & Paik, 2009; Perry, Sameulson, Malloy & Schiffer, 2010). Such findings are normally interpreted in line with the pervasive category-learning view of word learning, with variability "moving the focus from a specific value on a feature dimension to considering a range of acceptable values for category inclusion", "highlighting what features are *not* critical for a category" or "giv[ing] infants better representations of categorical boundaries" (all from Perry et al, 2010). But as we have already seen, models of word meaning based on feature-values for inclusion/exclusion or drawing categorical boundaries cannot work, as it is not possible to draw the boundaries in the right place.

An exemplar account easily accommodates such finding with no need for categories or boundaries. If a child has ten very similar exemplars of *dogs* (e.g., all golden retrievers) and then sees a new dog (e.g., a poodle), she will have no sufficiently similar exemplar stored, with its label, in memory. But suppose instead she has ten very different exemplars (e.g., a single labrador, beagle, bulldog, husky, daschund, terrier etc.). Individually, none is more similar to the poodle than is the Labrador, so a naïve single-nearest neighbour exemplar model will fail. But a *k*-nearest-neighbours model will succeed, because each of these dogs is similar to the poodle in a different way (e.g., colour, size, temperament), supporting the analogy of the label from these exemplars to the new one. As we saw above, this has long been understood in the general categorization literature, but these insights have not been incorporated into theories of word learning.

An exemplar account has another very important advantage in this scenario. The dimensions along which similarity is computed are not fixed, as they have to be under a prototype model (else there is no way to define the prototype). Rather, the dimensions along which similarity is computed – on the fly in real time – depend on the task in hand. In the context of a decontextualized laboratory-based word-learning task based on 2D images, the only dimensions on which information is available are visual ones such as size and colour. In the real world, people dealing with members of the category *dog* are trying to accomplish some task such as, for example, deciding whether to stay away from a particular dog for fear of being bitten. In this case, colour is irrelevant and temperament is everything. Clearly the best way to make a judgment about the likely behaviour of a novel dog is by analogy with particular previously-encountered dog exemplars that are similar on the relevant dimension – here temperament – not the average on this dimension of an idealized prototype. The fact that the average dog is easy going does not mean that this one will be. As we will see in more detail in the final section, we have no idea what properties will turn out to be relevant for some future goal. Consequently, we have no choice in storing all of the detail we can of particular exemplars. If we retain some – those that are "relevant" for defining the "category" (e.g., size, shape and colour) – and throw away others (e.g., temperament) we will be baffled when a new scenario arises.

2. Morphologically inflected words

English, as a language with relatively impoverished inflectional morphology, marks verbs for tense (e.g., *play**s*** [present] vs *play**ed*** [past]) and person/number (*She play**s*** [3rd person singular]/*They play* [3rd person plural]), and marks nouns for number (e.g., *one dog* [singular] *two dog**s*** [plural]). Many languages have much richer systems of inflectional morphology (e.g., noun and verb marking in languages such as Polish and Finnish, which we will consider shortly).

We begin, however, with a system that has attracted a great deal of research attention, due, in part, to its apparent simplicity: English past-tense marking. Setting aside irregular verbs (e.g., *sing/sang*), regular forms appear, at first blush, to be created by a "regular past-tense rule, which adds the suffix *-ed* to the end of a verb to indicate that the event referred to by the verb took place before the speech act (e.g. *walk-walked*)" (Prasada & Pinker, 1993: 2). This rule, which can be summarized informally as *[VERB]+ed*, is a classic example of the type of abstraction posited by abstraction-based theories; one that is "capable of operating on any verb, regardless of its sound" and that therefore "affords unlimited productivity" (Prasada & Pinker, 1993: 2). Although this particular formulation sits squarely within a *generativist-nativist* framework (and corresponds to a *rule-based* model of categorization), *usage based-constructivist* accounts also discuss "the formation of a…schema for regular inflection" (Maslen, Lieven, Theakston & Tomaselo, 2004: 1332-3), which would seem to correspond to a *prototype* model.

Whether one formulates *[VERB]+ed* as an abstract rule or an abstract prototype, it is again impossible to formulate the abstraction in such a way as to account for descriptive facts about the system (i.e., to rule in all, and only, possible forms), let alone empirical data from studies with adults and children. With regard to descriptive facts about the system, apparent "regular *-ed*" forms are not all created equal, but cluster into what Albright and Hayes (2003:127) call phonological "islands of reliability". For example, every English verb that ends in a voiceless fricative (*f, th*, *s* or *sh*) has a past-tense form ending in *-t* (e.g., *missed, hissed, wished*). Verbs that end in *-t* or *-d* (except irregulars) tend to have a past-tense form ending in *schwa+d* (e.g., *tended, needed, voted*), while those that end in *-b, -g* or *-n* tend to have a past-tense form ending simply in *-d* (e.g., *rubbed, sagged, planned*). This is clearly a problem for a *generativist-nativist* abstract rule "capable of operating on any verb, regardless of its sound" (Prasada & Pinker, 1993: 2), but it is equally problematic for a *usage based-constructivist* abstract "schema for regular inflection", since the system requires not one schema but at least three (we are back to the *lumping or splitting* problem raised in the previous section).

In fact, empirical data from acceptability judgment and production studies with adults and children demonstrate that we need neither one regular rule/schema, nor three, but none; i.e., an exemplar account. These studies (Albright & Hayes, 2003; Ambridge, 2010; Blything, Ambridge & Lieven, 2018) demonstrate that both the acceptability and production probability of "regular" past-tense forms for a given novel verb (e.g., *blafed, bredged, chooled, daped*) are predicted by its phonological distance from existing stored "regular" past-tense forms (e.g., *blafed* is similar to *chased* and *raced*; *bredged* to *wedged* and so on)[3]. Importantly, because phonological distance is graded and continuous, the finding of a correlation between acceptability/production probability and phonological distance to stored forms cannot be simulated by a realization rule that spells out regular "ed" as *-t, schwa+d* or *-d* in particular phonological contexts. Similar findings are observed in studies of over-regularization errors with English irregular verbs such as **sitted*, **goed* or **bringed* (e.g., Stemberger, 1993; Marchman, 1997; Marchmann, Wulfleck & Ellis Weismer, 1999 Hartshorne & Ullman, 2006 [though only for girls], Kidd & Lum, 2008). These findings suggest that, in both comprehension and production, children are retrieving individual exemplars of past-tense forms and using phonological analogy across them.

Given that abstraction-based accounts (from both theoretical perspectives) fail to account for one of the simplest apparent abstractions one could imagine (*[VERB]+ed*), we should not be surprised to learn that they fail for more complex systems, such as systems of verb and noun inflection in languages such as Polish, Finnish, Estonian, and Lithuanian. Indeed, for such systems, it is not clear what abstractions would be possible even in principle, given that the correct "ending[4]" varies not only across the verb person/number and noun case-marking paradigms, but according to properties such as gender, conjugation/declension class and phonological properties of the NOUN or VERB "stem". A lumping approach does not work, because a generalization such as *[NOUN]+[CASE MARKER]* is at far too high a level of abstraction to explain anything about the system (it is not the case that *any* individual case marking morpheme can be applied to *any* noun), and furthermore is inaccurate for systems that incorporate changes to the NOUN (or VERB) "stem". A splitting approach does not work because, as usual, once you've started splitting, you can't stop. Grammar books for such languages typically list around five conjugation/declension classes that capture broad generalizations, but split many of them into subclasses, some with just a handful of members. For example, Rasanen, Ambridge and Pine (2016) note that although the Finnish verb paradigm *lacks* conjugation classes *per se*, phenomena such as vowel insertion, vowel harmony and consonant gradation result in descriptive schemes of verb inflection that posit as many as 46 different phonologically-based classes.

We should therefore not be surprised to learn that studies of both verb and noun morphology (Rubino & Pine, 1998; Maratsos, 2000; Maslen et al., 2004; Aguado-Orea &

---

[3] In the studies of Albright and Hayes (2003) and Blything et al (2018), a simple exemplar-based model (a version of Nosfosky's, 1990, Generalized Context Model) was outperformed (at predicting participants' responses) by a multiple-rules model that posits a formal rule for each "island of reliability". However, Keuleers (2008) and Chandler (2010) argued that this difference was due to a number of simplifying assumptions adopted in Albright and Hayes' instantiation of the GCM, and show that two alternative exemplar-based models - the Tilburg Memory Based Learner and Skousen's (1989) Analogical Model – perform comparably with the multiple-rules model. Of course, achieving comparable performance is – on its own – no reason to favour these exemplar-based models over Albright and Hayes' (2003) multiple-rules model. However, this comparable performance becomes particularly impressive when it is borne in mind that the multiple-rules model was custom built for the particular case of English past-tense inflection, while the three analogical models are instantiations of much more general models of linguistic, and even nonlinguistic, categorization.

[4] In fact, some distinctions are marked additionally or solely by changes to the stem, such as vowel or consonant lengthening.

Pine, 2015; Leonard et al., 2002; Dąbrowska, 2004; Dąbrowska & Szczerbiński, 2006; Dąbrowska, 2008; Krajewski et al., 2011; Kunnari et al., 2011; Kirjavainen et al., 2012; Kjærbæk et al., 2014; Räsänen et al., 2016; Saviciute et al., 2017, Engelmann et al, submitted; Granlund et al, submitted) yield three findings that constitute evidence for an exemplar account. The first is an effect of token frequency of the individual target form: The greater the frequency of a particular "ready inflected" verb or noun form in the input, the greater the rate at which children produce it correctly, and the lower the error rate, in both naturalistic and experimental contexts. The second is an effect of phonological neighbourhood density: The greater the number of phonological "friends" or "neighbours" – forms that are phonologically similar to the target and that take the same inflectional ending – the greater the rate at which children produce the target form correctly, and the lower the error rate. The third is an effect of competition: When children do not produce the correct target form in an experimental study, they generally produce either a more frequent form of the target word (e.g., a 3pl verb form in place of 3sg; an accusative or genitive noun form in place of dative or instrumental), or overgeneralize a higher frequency ending from a different conjugation/declension class.

It is important to note that these frequency effects at the levels of both target and competing forms are not found only in very complex systems, where children have no "choice" but to store a multiplicity of individual exemplars. They are found also for systems that are virtually exceptionless, such as English 3sg *-s* marking (Rasanen, Ambridge & Pine, 2014) and Japanese past/nonpast marking (Tatsumi, Ambridge & Pine, 2018), where there is, no "need" to store individual ready-inflected forms (e.g, *fits, plays, runs, walks*) at all; since all could in principle be generated from the bare/nonfinite form (e.g., *fit, play, run, walk*). They are also found in studies of infants' production of monomorphemic single words, which require no "inflection" at all (e.g., Sosa & Stoel-Gammon, 2012).

These findings suggest that there really is no alternative to some form of exemplar account that posits no additional stored abstractions such as English past-tense *[VERB]ed* or plural *[NOUN]s*. The only debate that we should be having is exactly what type of exemplar account is correct. One possibility is that different tokens of a particular form (e.g., *plays*) – for example those produced by different speakers – are recognised as somehow "the same", and that each witnessed token increments the representational strength of that form in memory. This incrementation would need to occur according to some function that yields diminishing returns for each additional instance, in order to account for the fact that the relationship between frequency and performance is, roughly speaking, log-linear, rather than linear. Although this is probably the type of account that most researchers have in mind when they think of frequency-sensitive storage, it suffers from an important weakness: As we will see in Section 5, learners recognize different tokens of "the same" verb form (e.g., *plays*) produced by different speakers; a finding that is difficult to explain under the assumption that fine-grained phonological information is thrown away.

This type of finding suggests the need for a more radical exemplar account, under which each and every individual token is stored separately. As well as explaining effects such as speaker-identification, this account naturally explains frequency effects – at the level of both target and competing forms – without having to posit a notion of "representational strength" or a competition mechanism. On this view, token frequency effects are observed simply because the more slightly-different exemplars of a particular form are stored, the greater the likelihood of one being retrieved. To offer an analogy, the more slightly-different red marbles there are in a bag of mixed colour marbles, the easier it is to pull out a red marble when you need one (a token frequency effect at the level of the target form); and the less likely you are to erroneously pull out a blue one (a token frequency effect at the level of competitor forms). This account also naturally yields the diminishing returns that are a signature of frequency effects: Adding a red marble to a bag that contains 1 red marble and 99 blue marbles almost doubles the probability of retrieving a red marble. Adding a red marble to a bag that contains 50 red marbles and 99 blue marbles barely affects the odds at all.

3. *N*-grams

At the word level, bigrams are two-word sequences, trigrams are three-word sequences and, generally, *n*-grams are *n*-word sequences. For example, the utterance *You have another cookie right on the table* (spoken to Brown's, 1973, Eve in her first recording) includes the bigrams *you+have, have+another, another+cookie* (etc.), and the trigrams *you+have+another, another+cookie+right* (etc.). As we will see in more detail below, speakers show effects of *n*-gram frequency in production and comprehension. Yet, unlike the structures discussed in the other sections, mainstream abstraction-based accounts of child language acquisition – on both sides of the Chomskyan theoretical divide – include little-to-no role for *n*-grams, at least in the adult system.

Accounts in the generativist, Chomskyan tradition explicitly reject the notion of linguistic rules or representations based on "strings of words, rather than on their structural representations" or that "mention only linear relations" (Crain & Nakayama, 1987: 522). These accounts do, of course, allow for combinations of *syntactic constituents*, with the order specified by the setting of a head-direction parameter. For example, the DETERMINTER PHRASE (DP) *another cookie* would be formed by combining the DETERMINER *another* and the NOUN PHRASE *cookie* (e.g., Abney, 1987). But the bigram *another+cookie* would not itself be stored; let alone a bigram such as *cookie+right* that violates constituent structure.

Accounts in the constructivist tradition (e.g., Tomasello, 2003) do assume that children store strings such as *another+cookie* (though it is less clear whether they assume that children also store strings such as *cookie+right* that do not constitute a semantic, functional or communicative "unit"). But, they also assume that – with the exception of high-frequency frozen-phrases or idoms (e.g., *I+dunno*; Bybee & Scheibman, 1999) – stored strings are characteristic of an early rote-learned stage, and are largely replaced, in the adult system, by abstract patterns formed by analogy. For example, a learner might analogize across stored strings such as *the+cat, a+dog* and *another+cookie* to eventually form an abstract *[DETERMINER] [NOUN]* construction, which can be used to form any (semantically appropriate) combination; and that, therefore – in its mature adult state – is virtually indistinguishable from a generativist style DETERMINER PHRASE.

The claim that at least some of these early stored strings are replaced and effaced by these later abstractions is rarely made explicit. But it is not clear what else could be meant by the idea of linguistic "representations becoming fully abstract" (Dittmar, Abbot-Smith, Lieven & Tomasello, 2008: 581), of "the development of a more schematic and abstract inventory of conventionalised constructions" (Lieven, Salomo & Tomasello, 2009: 505), by the claim that children "construct…abstract syntactic representations in the course of development" (McClure, Pine & Lieven, 2006: 718), or that their "linguistic representations…become abstract and productive" (Savage, Lieven, Theakston & Tomasello, 2006: 29).

In sum, with one exception – Abbot Smith and Tomasello (2006), discussed in detail below – generativist and constructivist accounts alike assume that at least some adult knowledge of language – probably the majority – consists of knowledge of abstract rules, syntactic structures or constructions, rather than individual exemplars, strings or *n*-grams. Therefore, any evidence that children and adults show evidence of storage of individual *n*-gram exemplars would be problematic for all abstraction-based accounts of language acquisition. Indeed, *n*-grams constitute a particularly serious problem for these accounts. Unlike the abstract morphological and sentence-level constructions discussed in the previous and following sections, it is not clear what many abstract *n*-gram-level constructions would even look like. *Some* possible abstract two-element constructions would at least be meaningful in terms of the grammar (e.g., *[DETERMINER] [NOUN PHRASE]*). But what kind of abstract *n*-gram construction could encode knowledge of a bigram or trigram such as *cookie+right* or *cookie+right+on*? What does the idea of an "abstract *n*-gram construction" even mean? It is far from clear how stored strings such as *cookie+right* and *cookie+right+on* could be replaced even potentially by something more abstract.

Thus, any evidence that children and/or adults have knowledge of individual *n*-grams would be problematic for any non-exemplar account, whether constructivist or generativist. In fact, there is a wealth of such evidence from studies show faster processing and/or fewer production errors for higher than lower frequency *n*-grams for both adults (e.g., Liberman, 1963; Krug, 1998; Bybee & Scheibman, 1999; Jurafsky, Bell, Gregory, & Raymond, 2001; Sosa & MacFarlane, 2002; McDonald & Shillcock, 2003; Pluymaekers, Ernestus, & Baayen, 2005; Arnon & Snider, 2010; Tremblay & Baayen, 2010; Siyanova-Chanturia, Conklin, and van Heuven, 2011; Janssen & Barber, 2012) and children (Bannard & Matthews, 2008; Matthews & Bannard, 2010; Arnon & Clark, 2011). It is important to note that the *n*-grams in these studies were not part of high-frequency idioms or frozen phrases in the adult grammar (which have to be stored under any account); neither were they necessarily syntactic constituents or meaningful chunks (e.g., *got any; your truck; of milk*; from Bannard & Matthews, 2008)

The most straightforward way to account for these findings is simply to assume that learners are storing wholesale strings that they hear (e.g., *You have another cookie right on the table*), paired with their meanings; i.e., to assume an exemplar model[5]. Note that maintaining an inventory of the individual bigrams (*you+have, have+another, another+cookie, cookie+right, right+on, on+the, the+table*), trigrams (*you+have+another, another+cookie+right, right+on+the, on+the table*), four-grams (*you+have+another+cookie, cookie+right+on+the, right+on+the+table*), and so on would require not less storage, but many times more. And even if it were somehow possible to store *You have another cookie right on the table* as some form of abstraction – whatever that might mean – this abstraction would efface the *n*-gram frequency information that is evidenced in these studies.

One possible alternative model that could explain these findings is a "hybrid model comprising both abstractions and the retention of the exemplars of which those abstractions are composed" (Abbot-Smith & Tomasello, 2006: 276). But this risks giving us the worst of both worlds: The profligacy of an exemplar model – presumably the reason that most language acquisition theories favour abstraction-based accounts in the first place – plus the poor-data-coverage of a prototype model, coupled with more profligacy: The abstract prototype must be stored in addition to the exemplars, even though it adds no explanatory power.

---

[5] Baayen, Hendrix and Ramscar (2013) describe a naïve-discriminative-learning model that yields four-word *n*-gram effects without storing either sentence-level lexical strings (e.g., *A British provincial city*) or their component *n*-grams. However, because the model relies on a richly structured – albeit unordered – set of meaning representations (e.g., {A, BRITAIN, ISH, PROVINCE, IAL, CITY}), it is unclear whether it is really eschewing exemplar storage altogether, or instantiating it at the level of stored semantic representations. Furthermore, eschewing storage at the lexical-phonological level is problematic, given findings (summarized in Section 5) that listeners store fine-grained phonological representations of the utterances that they hear, rather than merely extracting semantic representations.

All of the arguments and findings set out above also hold at the syllable level, where learners demonstrate sensitivity to the transitional probabilities of individual syllables (e.g., *pre+tty+ba+by*; Saffran, Aslin & Newport, 1996; Aylett & Turk, 2006). They also apply at the level of individual phonemes (e.g., learners of English know that *f+t* is more common than *v+t*; Mattys & Juscyzk, 2001). The only way to explain these findings would be to assume that, in addition to word-level *n*-grams (*pretty+baby*), learners are also storing syllable-level and phoneme-level *n*-grams (and, under a hybrid account, more abstractions too). Again, this would require not *less* storage than simply storing whole utterances, but many times *more*.

4. Sentence level constructions

While they disagree with regard to the timing and technical details, if there is one thing on which all mainstream accounts of child language acquisition agree, it is that by – at the latest – around 3;0 (Tomasello, 2000; 2003) children have abstract knowledge of word order. Constructivist, early-abstraction and generativist-nativist accounts claim, respectively, that learners of English have acquired "some kind of abstract, verb-general, SVO transitive construction" (Tomasello, 2000: 216), "detected the abstract word-order pattern of English transitive sentences" (Gertner et al, 2006: 686) and "set parameters correctly at an extremely early age…includ[ing] word order" (Wexler, 1998: 29). Indeed, in many of my previous writings, I have been no exception, concluding for example (in Ambridge & Lieven, 2011: 239) that "by 2;0, children almost certainly have at least some[6] abstract, verb-general knowledge of the basic word-order rules of English (i.e. the SUBJECT VERB OBJECT transitive construction)".

Yet, once again, the notion of an abstract SVO construction, schema or rule falls at the first hurdle of accurately describing the adult grammar, before we even consider data from empirical studies. It is transparently not the case that any verb can appear in the VERB position (e.g., **The comedian* ***laughed*** *her* [c.f., *The comedian made her laugh*]), nor any OBJECT in the OBJECT position (c.f., *The comedian laughed* ***a hearty laugh***), nor any SUBJECT in the SUBJECT position (c.f. ****The magician*** *disappeared the rabbit vs* ***Stalin*** *disappeared his enemies*). We cannot solve the problem by simply appending labels such as "intransitive" or "transitive" to verbs' lexical entries, because the acceptability of such sentences depends on the other arguments, and – in adult judgment studies – is finely graded (Ambridge, Pine, Rowland & Young, 2008; Ambridge, Pine, Rowland, Jones & Clark, 2009; Bidgood, Pine, Rowland & Ambridge, submitted). The dilemma again is one of lumping versus splitting; and the notion of a single abstract SVO transitive construction is perhaps the most extreme and inaccurate example of lumping that we have encountered so far. Consider the various different sentence types that must be subsumed by a unitary SVO transitive construction (adapted from Ambridge & Lieven, 2015):

Contact (non-causative)	[AGENT] [ACTION] [PATIENT]	John hit Bill
Causative	[CAUSER] [ACTION] [UNDERGOER]	John broke the plate
Experiencer-Theme	[EXPERIENCER] [EXPERIENCE] [THEME]	John feared Bill
Theme-Experiencer	[THEME] [EXPERIENCE] [EXPERIENCER]	John frightened Bill
"Weigh" Construction	[THING] [MEASURE/COST/WEIGH] [AMOUNT]	John weighed 100lbs
"Contain" Construction	[CONTAINER] [CONTAIN] [CONTENTS]	The tent sleeps four people

[6] Though, as hinted at by the caveat "at least some", I was already beginning to doubt the meaningfulness of the concept of a standalone abstract SVO construction

What is particularly problematic in this case is that certain concrete instantiations of this would-be unitary construction are not just radically different, but polar opposites (e.g., *John feared Bill* vs *John frightened Bill*). Semantically, a putative SUBJECT prototype would have to encompass both a frightener and one who is frightened; a chaser and a fleer; a giver and a receiver. As we saw in Section 1, with the example of large (wooden) and small (metal) spoons, a prototype category structure cannot represent these types of non-linearly-separable distinctions. If a *frightener* is a prototypical SUBJECT, then a *frightenee* is as non-prototypical a SUBJECT as one can imagine. Formal linguists (e.g., Newmeyer, 2003) may object that syntactic subjecthood is entirely independent of semantics; but as well as being empirically false (as noted above, participants rate some transitive subjects as better than others), this objection misses the point. The very function of word-order in morphologically impoverished languages such as English is to convey semantics (c.f., *The dog bit the man; The man bit the dog*). If those semantics can flip entirely depending on the identity of the verb (*frighten/fear, chase/flee, give/receive*), then what is conveying the meaning simply cannot be an abstraction that is insensitive to the identity of the verb (i.e., some type of SVO construction, rule or schema).

The only solution is to admit the identity of the verb into the representation; but this is to jump out of the frying pan of lumping, and into the fire of splitting. Suppose that we posit a separate construction for each of the six types outlined above. This does not solve the problem because, to take just one example, the [EXPERIENCER] [EXPERIENCE] [THEME] construction inappropriately lumps across a variety of events (e.g., *John heard / saw / spotted / noticed / recognized Bill*) that differ considerably with regard to the nature of the interaction (who is the one "doing something" in each of these cases? Sometimes John, sometimes Bill). Indeed, individual instantiations of the (putative) [EXPERIENCER] [EXPERIENCE] [THEME] construction vary continuously in their grammatical acceptability, speed of processing and production probability (Ambridge, Bidgood, Pine, Rowland & Freudenthal, 2016; Bidgood, Pine, Rowland & Ambridge, submitted), who also find the same for the (putative) [THEME] [EXPERIENCE] [EXPERIENCER] and [AGENT] [ACTION] [PATIENT] constructions, as well as their passive equivalents.

Again, then, we would expect the experimental data to be problematic for the idea of an abstract representation of word order, whether a generativist-nativist style formal rule that is insensitive to the identity of the verb and its arguments, or a constructivist style SVO construction with prototype structure. And this is exactly what we find. Even though virtually all child studies have given abstraction-based theories a head start by equating SVO with [AGENT] [ACTION] [PATIENT] or [CAUSER] [ACTION] [UNDERGOER] – ignoring all the other possible types listed above – their findings are covered with the fingerprints of exemplar storage.

Ambridge & Lieven (2011:221) summarized 14 elicited-production studies in which novel verbs were elicited in an SVO transitive construction (e.g., *He's tamming it*), having been presented solely in non-transitive forms during training. Across these studies, the majority of arguments – particularly for the younger children – were pronouns (e.g., *He's tamming it*) (e.g., 90% in Dodson & Tomasello, 1998). On its own, this finding could simply reflect discourse tendencies of English. However, Childers and Tomasello (2001) found that training children on overlapping exemplars with English verbs (e.g., *He's pushing it*) increased the proportion of two-year-olds who produced an SVO utterance with novel verbs (e.g., *He's tamming it*) from 45% to 85%. Similarly, Akhtar's (1999) weird word order study found that children used pronouns for around 50% of all arguments when producing SVO transitives (e.g., *He's tamming it*), but *never* when imitating a weird word order produced by the experimenter (e.g., **Elmo the car gopping* not *He it gopping*) (see also Abbot-Smith, Lieven & Tomasello, 2001; Matthews, Lieven, Theakston & Tomasello, 2004; 2007; Savage, Lieven, Theakston & Tomasello, 2003). This advantage for pronoun-based over full-noun based SVO transitives (e.g., *He's meeking it* vs *The dog's meeking the car*) is also seen in comprehension studies (e.g., Childers & Tomasello, 2001, Study 2).

These findings (which constitute clear evidence against generativist-nativist style context-free word order rules) are usually taken as evidence for constructivist style *slot-and-frame patterns* or *lexically-based schemas* (e.g., *He's ACTIONing it*). The question is whether this is simply a metaphor for the kinds of on-the-fly analogical generalizations posited by exemplar accounts, or whether such abstractions are somehow stored and represented independently. Most constructivist accounts do not address this question directly, but seem to at least hint implicitly at the latter (e.g., Chandler, 2010, attributes such a position to Croft & Cruse, 2004; Goldberg, 2006; and Langacker, 2009). Indeed, for my own part, I have often written about children acquiring slot-and-frame patterns without stopping to think which of the two possibilities I intend or imply. As we have already seen, the paper by Abbot-Smith and Tomasello (2006: 281-282) is an exception in explicitly advocating for the latter view [emphasis added]:

> In view of the prevalence of item-based effects and frequency effects in syntactic acquisition (and which remain to some degree in adult language usage, see Dabrowska 2004), exemplar models of categorization are more attractive than a 'pure' prototype-abstraction model in which the extraneous details of original instances are completely lost. Furthermore, such exemplar-learning models are perhaps better able than pure prototype models to explain patterns of family resemblance in syntactic and morphological categories where there is no central tendency (e.g., Bybee 1995). However, we would question the assumption that more abstract prototype categories are only generalized online and leave no permanent representational change. Even in exemplar models every time an exemplar is comprehended, its representation must change in some way, even if this merely involves registering frequency. If the comprehension or production of a novel utterance involves 'summing over' similar sets of exemplars, the frequency with which a set is called upon probably also leaves a trace. **Therefore, if the mutual similarities of a particular collection of exemplars (such as transitive sentences) are 'summed over' regularly, we believe this is highly likely to permanently change the user's linguistic representations in some way equivalent to the formation of some kind of more abstract representation**. A resolution to the drawbacks of both 'pure' prototype and 'pure' exemplar learning models is **a hybrid in which much of the extraneous details of original instances are retained but where some kind of more abstract schema is gradually formed on the basis of these**.

As Abbot-Smith and Tomasello (2006: 282) themselves note "It may of course prove difficult to empirically differentiate such a hybrid model from certain 'pure' exemplar-learning models". Here, then, are three non-empirical, theoretical arguments for a pure exemplar-learning model. The first is simply Occam's Razor. If we are positing exemplar storage anyway, we should not posit some additional abstractions on top unless they add explanatory power; and to our knowledge, neither Abbot-Smith and Tomasello (2006) nor any other paper has made the case for a phenomenon that cannot be captured by a pure exemplar model, whether in the domain of language, or of learning and categorization more generally.

The second argument is the familiar lumping or splitting problem. We have already seen that a high-level abstraction such as SVO cannot accommodate differences between *frighten/chase/give* type sentences and *fear/flee/receive* type sentences (or, indeed, verb-by-verb differences within these two classes). The same is true for lower-level abstractions such as the putative slot-and-frame patterns *He's [ACTION]ing him/it*, since utterances constructed using such a template can again have opposite meanings (e.g., *He's chasing it* vs. *He's fleeing it*). In general, as we have seen in all the domains covered in the present article, it is never possible to posit exactly the right abstractions; those that rule in all possible sentences and rule out all impossible ones.

The third reason to favour a pure exemplar-learning based account over the hybrid account posited by Abbot-Smith and Tomasello (2006) is that the additional abstractions posited raise more questions than they answer (if, as per the first point above, they answer any). Which slot-and-frame patterns do learners abstract, and why these ones (token frequency? type frequency? communicative function)? How do learners move from semi-abstract slot-and-frame patterns (e.g., *He's [ACTION]ing it?*) to a fully abstract SVO construction? Are the various abstractions assumed independent or linked? For example, is the SUBJECT category in the fully abstract adult transitive construction (SVO; She's dancing) the same or different to the SUBJECT category in the intransitive construction (SV; e.g, *She danced*), the dative construction (SVOO; *She gave him a book*), and the passive construction (SV by O; *She was chased by him*). If they are "linked", what exactly does this mean in terms of representation and processing? What about when these constructions are combined (e.g., the dative and passive, to yield SVO by O; *She was given a book by him*). And are the answers to these questions the same or different when we are talking about lower-level, less-abstract slot-and-frame patterns (e.g., *She's [ACTION]ing it; She [ACTION]ed; She gave [PERSON] [THING]; It got [ACTION]ed by it* etc.)?

A pure exemplar account bypasses all of these difficulties. A novel utterance is produced by analogy with all the stored exemplars that are sufficiently close to the target meaning. This might be an entire stored utterance (e.g., *She's dancing*), a set of exemplars with high semantic overlap (e.g., *She's running, She's jumping, She's dancing*; equivalent to the slot-and-frame metaphor) or – failing both of these options – a set of exemplars with lower, but still sufficient, overlap (e.g., *The girl danced; Sue danced; Jim is dancing; He's dancing*). Note that analogy across what would normally be considered "different constructions" provides a ready-made explanation of *construction conspiracy* phenomena (Abbot-Smith & Behrens, 2006); for example, the finding that the acquisition of the German *sein* passive (e.g., *Der Reis war gekocht,* 'the rice was cooked') is boosted by experience with the *sein* copula construction (e.g., *Der Reis war Schwarz*; 'the rice was black').

Of course, the devil is in the detail, in the need to explain the basis for analogy; but the exemplars+abstractions account is not immune here, as it must similarly explain the basis for analogy that leads to the stored slot-and-frame patterns and higher-level abstractions. If the basis for analogy can be figured out, there is no need to posit the additional stored abstractions: The same analogies can be used by a pure exemplar model to generate utterances as they are needed, bypassing altogether problems regarding links between constructions, combining constructions and so on.

Finally, what about syntactic priming studies? Doesn't "priming evidence support the existence of abstract syntactic representations" (Branigan & Pickering, 2017: 9)? Syntactic priming is a phenomenon whereby if presented with one exemplar of a particular syntactic construction (e.g., *The vase was broken by the hammer*), speakers are more likely than they would otherwise have been to re-use the same construction when asked to describe a subsequent picture or animation (e.g., to say *The bricks were pushed by the digger*, rather than *The digger pushed the bricks*). The claim is that syntactic priming effects of this type reflect "repetition of aspects of abstract linguistic structure" (Branigan & Picking, 2017: 2); i.e., that a stored abstract representation of (for this example) the passive construction is activated by the prime, and then used by the speaker to produce her own utterance.

The existence of syntactic priming effects for both adults and children is uncontroversial (see Mahowald, James, Futrell & Gibson, 2016; and Branigan & Pickering, 2017, for comprehensive and up-to-date reviews). What is controversial is the claim that the phenomenon "is best characterized in terms of the presence or absence of (re)activation of purely abstract syntactic representations" (Feldman & Milin, 2017: 22; see other responses to Branigan and Pickering's target article by Kempsona & Gregoromichelaki, 2017, and O'Grady, 2017, as well as rival accounts based on the notion of implicit learning; e.g., Chang, Dell, Bock & Griffin, 2000).

Indeed, the exemplar approach advocated in the present article can, in principle, easily accommodate syntactic priming effects (see also Goldwater, Tomlinson, Echols & Love, 2011, for a different but compatible account of syntactic priming as exemplar-driven analogy): A prime sentence such as *The vase was broken by the hammer* activates in memory not an abstract representation of the passive, but stored concrete exemplars that meet some threshold for both surface similarity (phonological and suprasegmental) and semantic similarity, both passive (e.g., *The man got run over by a bus; The window was smashed by a ball*) and active (e.g., *The man broke the vase*). When generating a subsequent utterance, the speaker does so by means of on-the-fly analogy across relevant stored examples, exactly as she does when generating *any* utterance. The only difference is that a number of concrete passive exemplars, having been recently activated by the prime, are more available for retrieval than they would otherwise have been.

An exemplar account of priming can explain a number of findings that are problematic for abstraction-based accounts. First, priming effects are small and probabilistic. A recent meta-analysis (Mahowald, James, Futrell & Gibson, 2016) found that, on average, a construction that occurs 50% of the time without priming occurs 63% of the time if it is preceded by a form that uses the same syntactic structure as the target (but shares no words in common; i.e., no lexical overlap). Thus, when primed with a passive construction (which, without priming has a production probability of well under 50%), a speaker is still overwhelmingly more likely to produce an active than a passive. Indeed, in one recent priming study (Bidgood et al, submitted), almost half of the 4-6-year-old children tested (25/60) produced no passives at all, despite each being primed with a passive on 18 trials. Such findings are expected on an exemplar account, given that a passive prime (e.g., *The vase was broken by the hammer*) activates both passive *and active* forms that meet some threshold for similarity, and that active exemplars are – on one estimate – around 80 times more frequent than passive exemplars in spontaneous speech (Ambridge, Bidgood, Twomey, Pine, Rowland & Freudenthal, 2015). Such findings are problematic for an account based on the notion of activating abstract representations, since, for cases such as passives, the most recently activated abstraction remains unlikely to be used.

Second, a very reliable finding is that the priming effect is boosted – the meta-analytic average of 63% production of the primed construction jumps to 77% – if lexical material is shared between the prime and the target (e.g., between ***The** vase **was broken by the** ball* and ***The** window **was broken by the** hammer*, or even – as in Savage et al, 2003 – between ***It got** pushed **by it*** and ***It got** broken **by it***). Abstraction-based accounts can explain this "lexical boost" only by positing some form of add-on, such as "a representation that encodes a binding between constituent structure and the lemma (syntactic component) of the lexical entry for the head" (Branigan & Pickering, 2017:12) or "short-term activation of explicit memory traces" (Rowland et al, 2012: 53). Such add-ons have no independent motivation; they are posited solely to explain the lexical boost. In contrast, an exemplar account explains the lexical boost naturally, with no need for additional assumptions: The prime preferentially activates stored exemplars with which it shares surface (lexical, phonological, suprasegmental) and semantic overlap, with these exemplars therefore more available for use

– via analogy – in subsequent production. An exemplar account also explains why the lexical boost appears to increase with development (Rowland, Chang, Ambridge, Pine & Lieven, 2012): The more exemplars have been stored in memory, the greater the probability of lexical overlap between the prime and a stored exemplar.

Third, an exemplar account naturally explains why priming is seen in cases where there is substantial lexical overlap between the prime and the target, but only debatable overlap in terms of abstract syntax. Bock and Loebell (1990) famously showed that passive sentences such as *The construction worker was hit by the bulldozer* prime intransitive locative (i.e., *non*-passive) sentences such as *The 747 was landing by the airport's control tower*. This is entirely expected on an exemplar account, on the basis of overlap between these utterances in terms of overlap in lexical/phonological material (e.g., Ziegler, Goldberg & Snedeker, 2018) and suprasegmental stress patterns. It is entirely unexpected on a standard linguistic analysis which views these two abstract constructions as having very different underlying syntactic structure. Again, then, the abstraction-based approach favoured by, amongst others, Branigan and Pickering (2017: 9) requires a special workaround; here the assumption that "abstract syntactic representations….are shallow and monostratal" (i.e., [NOUN PHRASE] [VERB PHRASE] by [NOUN PHRASE], as opposed to [SUBJECT] BE [VERB] by [OBJECT]). This solution is unsatisfactory, as it fails to account for the important semantic differences between the two constructions. For example, in *The worker was hit by the boss* (passive) and *The worker was sitting by the tree* (intransitive), 'the worker' is playing very different semantic roles (PATIENT and ACTOR respectively), which is precisely why traditional syntactic theories posit different syntactic structures in the first place. Even with this unsatisfactory workaround, the abstraction-based approach does not accommodate the findings of a recent modified replication of Bock and Lobell (1990), which suggest that this passive priming effect was driven solely by the lexical item *by* (Ziegler, Goldberg & Snedeker, 2018). Ziegler et al found that *by* was *both necessary and sufficient* for priming passive sentences: No priming of passives occurred following locatives that lacked *by* (e.g., *The 747 was landing next to* [c.f. *by*] *the airport's control tower*). Conversely, priming of passives *did* occur follow *active* locative sentences with by (e.g., *The pilot landed the 747 by the control tower*).

Fourth, an exemplar account naturally explains why priming is boosted if the prime and target share not only lexical overlap, but also overlap in the ordering of their semantic roles. Chang, Bock and Goldberg (2003) found that if syntactic structure (in the monostratal sense) is held constant, priming is boosted if the prime and target overlap in terms of their semantic roles. For example, a target sentence such as *The farmer heaped straw* [THEME] *onto the wagon* [LOCATION] is better primed by another THEME-LOCATION sentence (e.g., *The maid rubbed polish onto the table*) than by a similar LOCATION-THEME sentence (e.g., *The maid rubbed the table with polish*). Similarly, Ziegler and Snedeker (2018) observed priming at the level of the ordering of semantic roles, even without syntactic or lexical overlap. Again, this is entirely expected on an exemplar account under which the prime preferentially activates stored exemplars in proportion to the degree of both phonological and semantic overlap. And, again, it is highly problematic for an abstraction-based account. Branigan and Pickering (2017: 10) suggest that these findings "could have reflected a tendency to repeatedly assign thematic roles (e.g., Location) to grammatical functions (e.g., direct object) or to word-order positions (e.g., immediately following the verb)". But this tendency is not incorporated into their theoretical model; and, indeed, overlap in terms of particular semantics and particular word-order positions is more characteristic of an exemplar account than one based solely on overlap of abstract syntactic structure.

5. Phonetics and Phonology

For many speakers of British English, the *t* sound in words such as *Saturday* can be pronounced as either *t* (/t/ in IPA) or as a glottal stop (/ʔ/ in IPA), roughly *Sa'urday*, with the former used in more formal contexts. In many working-class accents, such as Cockney and Estuary English, the latter form predominates; to the extent that high-ranking British politicians who speak with received pronunciation sometimes – to widespread mockery – affect a glottal stop when trying to appeal to down-to-earth voters, for example when touring a factory (see also Dick Van Dyke's Cockney chimney sweep in *Mary Poppins*). These types of sociolinguistic effects are uncontroversial, and have been well known since at least Labov's (1963) famous Martha's Vineyard study (see Hay, 2018, for a recent review), and are perhaps why exemplar accounts have had more influence in phonetics and phonology than in other areas of language acquisition. Yet these effects are neglected entirely by mainstream accounts of word learning (Section 1), inflectional morphology (Section 2) and syntax (Section 4), which start from the assumption that learners represent an idealized word form (e.g., *Saturday*), from which the phonetic details of individual exemplars have been abstracted away. British speakers who switch between *Saturday* and *Sa'urday* depending on the sociolinguistic context therefore present a familiar dilemma for non-exemplar accounts: lump or split.

The lumping approach posits that learners have a single prototype representation of each word (in this case, a form somehow in between *Saturday* and *Sa'urday*), and some kind of high-level model which specifies how each of these idealized forms is phonetically realized in particular cases. This claim rapidly starts to look implausible, when it is borne in mind that learners would need a different realization model for every different type of speaker and accent they can recognize (e.g., older vs younger, working- vs middle-class vs upper-class; male vs female; British vs American English; Scottish vs English; Glasgow vs Edinburgh; Manchester vs Liverpool).

The splitting approach posits that, for each word, learners have a number of *different* prototypes for each word (e.g., *Saturday* and *Sa'urday*). But, again, once you start splitting, you can't stop. These two prototypes are certainly not sufficient to explain all of the possible pronunciations of *Saturday* that a British speaker might encounter and recognize as characteristic of a particular group, context or person. Are we then to posit separate *Saturday* prototypes for (a) older working-class women from Glasgow speaking in a formal context, (b) middle-class teenage boys from Manchester chatting informally, and so on all the way down to the level of individual speakers who we can recognize by their voices?

The exemplar approach sidesteps the lumping-or-splitting dilemma. Every token heard is simply stored as an episodic memory that contains both phonetic detail and factors such as the speaker identity and context. If the assumption of such rampant storage seems implausible, then consider the alternative: Is it *more* plausible that we represent and store a realization model for every combination of speaker and context that we encounter, and update these models in real time? And, if not, then what – other than an exemplar model – is the alternative?

As for the other domains of acquisition considered here, an exemplar account is again well supported by empirical data. Over 40 years ago, Craik and Kirsner (1974) found that adult listeners showed better recognition of words presented in the same than in different voices (male vs. female) at training and test. Palmeri, Goldinger and Pisoni (1993) found that this effect was undiminished when as many as 20 different voices were used, or when the gender of the speaker was held constant. Goldinger (1996) further demonstrated that, when *different* voices are heard at training and test, listeners show an advantage for *similar* over dissimilar voices, exactly as would be predicted under an exemplar model which assumes that learners are averaging on the fly across multiple stored exemplars (but not a simple nearest-neighbour exemplar model). Both the same- and similar-voice advantage held when training and test were separated by one day, but not one week, indicating that exemplar storage does not, of course, entail perfect memory; a point to which we will return in the final section.

Similar findings have also been observed for infants. Houston and Jusczyk (2000) found that 7.5-month olds showed a listening preference for previously-familiarized words, but only when the speaker was of the same gender at training and test. Although this effect had disappeared by 10.5 months, the adult findings above suggest older infants are not incapable of detecting speaker differences, but that their various stored exemplars of (say) *cup* are sufficiently similar to support recognition (which must be the case for adult speakers, since we are clearly capable of recognizing a familiar word produced by a new speaker). Neither are the exemplars rapidly lost. Houston and Jusczyk (2003) found that 7.5-month olds' preference for previously-familiarized words held a day later, but *only* when the speaker was the same at training and test.

As a consequence of these well-known and long-standing findings, exemplar accounts have become relatively well established in phonetics and phonology (e.g., Hay & Bresnan, 2006). Unlike for the other domains of language acquisition reviewed in the present article, there exist several exemplar-based (sometimes called "episodic") computational models that can simulate findings such as the speaker effects outlined above. These include Juscyzk's WRAPSA (1997), Goldinger's (1998) MINERVA, Morgan, Singh, Bortfeld, Rathbun and White's (2001) DRIBBLER, and Johnson's (2007) XMOD. In contrast, abstract-prototype models such as Kuhl's (1991) perceptual magnet approach (see Guenther & Gjaja, 1996, for a computational implementation), as well as rule-based models such as generative phonology (Chomsky & Halle, 1968) and Optimality Theory (Prince & Smolensky, 2008) cannot simulate these types of well-documented effects, because they operate at a level of abstraction that is far removed from individual exemplars. In summary, the advantages of exemplar models are generally well understood in phonetics and phonology, and it is time that other domains of language acquisition follow this lead.

6. Conclusion

In summary, in the domains of (1) word meanings, (2) morphologically inflected single words, (3) *n*-grams, (4) sentence-level constructions and (5) phonetics and phonology, abstraction-based accounts fail for two reasons. First, it is not possible to posit abstractions that delineate possible and impossible form; e.g., that (1) rule in pool *tables* and data *tables*, but rule out chairs, (2) rule in "mist" and rule out "missid" as the phonological form of the past tense of *miss*, (3) rule in the bigram *f+t* but rule out (probabilistically) *v+t*, (4) rule in both *John feared Bill* and *John frightened Bill* but rule out **John laughed Bill*, (5) rule in Speaker A but rule out Speaker B as the person who produced a particular word (e.g., *Sa'urday*). Second, for each domain, empirical data provide evidence of exemplar storage that cannot be captured by putative abstractions: e.g. speakers prefer and/or show an advantage for (1) exemplar variation even within word-meaning "category boundaries", (2) novel inflected forms that are similar to existing exemplars, (3) *n*-grams that have occurred frequently in the input, (4) SVO sentences with *he* as SUBJECT and *it* as OBJECT and (5) repeated productions of "the same" word that are phonologically similar or, better still, identical.

An exemplar account not only (1) avoids the problems facing abstraction-based accounts (all some versions of the "lumping or splitting?" problem) and (2) explains the exemplar effects observed in every domain but (3) provides a unifying account of language acquisition *across* all of these domains. Whether we are looking at the level of individual phonemes, single words, multi-word phrases or full sentences, an exemplar accounts offers exactly the same explanation: storage, and subsequent retrieval with analogy, of episodic memory traces including the phonetic detail of what was heard, the speaker's intended meaning, socio-cultural details (e.g., the setting, others present), and so on. In contrast, abstraction-based accounts offer an explanation of a simplified, idealized version of a phenomenon at one particular level, leaving crucial details to the morphologists, the syntacticians, the phonologists and so on. For example, an account of (1) word meaning might ask how children learn the action label *run*, ignoring (2) different morphological forms (*run+ran*), (3) *n*-gram phrases (*run+a+bath, run+for+office*), (4) sentence level combinatorial properties (*The coach ran the race* vs **The coach ran the sprinter*) and (5) phonological realizations produced by different speakers.

The fact that an exemplar account gives a unifying explanation of language acquisition across all domains is advantageous not solely for reasons of parsimony; the account also offers a ready-made explanation for correlations that are observed across different acquisition domains. For example, a number of studies (see Dabrowska, 2018, for an up-to-date review and novel adult data) have found correlations across learners on measures of vocabulary and morphosyntactic knowledge, even when looking at adults. This finding does not sit comfortably with generativist accounts under which the lexicon and the grammar are separate modules (e.g. Pinker's, 1999, "words and rules"), or with constructivist accounts under which speakers form a morphosyntactic abstraction when they have reached a critical mass of lexical exemplars (e.g., Marchman & Bates, 1994); particularly given that the finding holds for adults, who presumably reached any would-be critical-mass long ago. However, they are expected under an exemplar account, since exactly the same processes – storing episodic memories and generalizing across them in comprehension/production – are operational across domains and across the lifespan. The same point can be made for the observed bidirectional relationship between phonological and lexical development (see Stoel-Gammon, 2011, for a review). Although few researchers would find this relationship surprising, the traditional division of theoretical labour means that accounts of phonological and lexical development rarely make reference to one another. Again, an exemplar approach offers a unifying account of acquisition across these domains, and explains the observed correlations as reflecting the same learning process. More generally, an exemplar approach unites language acquisition with the acquisition of other cognitive skills (e.g., Chater & Christiansen, in press) such as nonlinguistic categorization (e.g., Mack et al, 2013) and recognition (e.g., Nosofsky, 2016). As we have seen, there even exists at least one proposal for how a generalized exemplar theory could be implemented in the brain, in terms of "connectivity between striatal neurons and neurons in sensory association cortex" (Ashby & Rosedahl, 2017: 472).

Despite the theoretical and empirical advantages of exemplar models, they have made very little headway in theories of language acquisition. Partly this is because many theorists seem not to have considered in any detail the question of representation, and exactly what that abstractions that they informally posit and discuss would entail. But when exemplar models are raised, many theorists quickly dismiss them as intuitively implausible. We therefore end by considering some of the most commonly raised objections to exemplar models.

Objection: It's not plausible that we store all this information

This is perhaps the major sticking point for exemplar models. The idea that we store in memory every linguistic episode that we have encountered (and, for that matter, every nonlinguistic episode too) is intuitively implausible. Partly this is because it doesn't feel that way subjectively (i.e., we cannot recall memories at will); partly because we are used to thinking about memory using the finite-storage metaphor of our times (a store house, a tape recorder, computer memory). But experiments with human memory show that we can recognize (and therefore must have stored) experiences that are not available via free recall, and challenge the notion of finite storage. Standing (1973) found that participants who were shown 10,000 photographs could, two days later, subsequently categorize a randomly-chosen subset of 160 as seen or unseen with an average success rate of 83%. And – aside from forgetting (discussed below) – the fact that this figure is not even higher is due not to too little exemplar storage, but too much: Difficulties in retrieval are caused in part by interference between similar memories; and with 10,000 images, a relatively high degree of overlap is inevitable (in addition to overlap of contextual details such as the testing room). Tellingly, pictures chosen to be particularly vivid (e.g., a dog with a pipe in its mouth) showed better recognition than normal pictures (e.g., a dog without a pipe), precisely because interference is minimized.

A second counterargument to the implausibility of such rampant storage is simply that the psycholinguistic data provide compelling evidence for it, in the form of – to pick perhaps the clearest example of those we have reviewed – frequency effects for particular *n*-grams (e.g., *drink+of+milk* > *drink+of+tea*). When compelling empirical data contradict our deeply-held intuitions, it is time to revise those intuitions. It is not intuitively plausible that the earth is round and revolves around the sun, that diseases are spread by invisible microscopic agents, that (almost) all cells contain a full set of instructions for building a human, or that we share a common ancestor with bananas. But we have learned to accept all of these facts because the data strongly suggest them.

A third and related counterargument is to ask what alternative explanation would be more plausible. For example, as we saw in the previous section, it is not clear how the ability to recognize individual accents and speakers can be explained by an exemplar model, other than by assuming that we maintain, for each speaker we can recognize, a realization model that transforms highly abstract phonological representations into the precise surface forms heard. This is not impossible, of course. But it is difficult to make the case that it is more intuitively plausible than the exemplar-storage alternative.

Objection: But we do forget things, right?

Yes. Exemplar models assume that everything is stored, but not perfectly, and not indefinitely. Of course, some details (those to which we are not attending when each episodic trace is created) are not stored in the first place. Subsequently, forgetting (or trace decay) happens at both the level of whole exemplars and of individual details within exemplars. Indeed, all of the major computational exemplar models reviewed above incorporate some kind of forgetting function; not merely for reasons of face validity, but because it is important for simulating human performance. For example, Hintzman's MINERVA2 simulation of human performance in dot-pattern classification tasks incorporates a bout of forgetting in which 75% of features (i.e., details within individual exemplars) are set to zero. Forgetting, of course, occurs in linguistic tasks too. As we noted above, the recognition advantage for same- and similar-voice exemplars observed by Goldinger (1996) is present after a day, but has disappeared by a week.

Objection: Aha! Isn't this "abstraction"?

Not necessarily. Suppose that, after a delay of one week, a participant can recognise *cup* as a previously-presented word, but shows no same-/similar-speaker advantage; i.e., no retention of phonetic detail. One possible explanation is that she has somehow merged the *cup* that she previously heard into idealized representations of *cup* at both the phonological/phonetic level (see Section 5) and the semantic level (see Section 1). However, this explanation is problematic for reasons that we have already explored in detail. It is not possible to posit an idealized phonetic representation of *cup* that rules in all – and only all – possible pronunciations, or an idealized semantic representation that rules in all – and only all – possible entities that might be referred to as a *cup* (including, for example, the World Cup, which is either a tournament or a statuette). In short, there is no plausible idealized *cup* into which this particular exemplar can be merged. The exemplar account offers a much more straightforward explanation: The participant has simply retained some details of the episodic memory (e.g., the semantics invoked for her by *cup* during the training session), but not others (i.e., fine-grained phonetic details).

The weakness of the abstraction-as-forgetting account becomes even more apparent when we consider other domains. Take, for example, the finding discussed in Section 2 (inflectional morphology) that, when asked to judge novel English past tense forms, participants show sensitivity to the fact that verbs ending *f, th, s* or *sh*, form their past tense with *-t*. It would seem very odd to argue that English speakers witness forms like *missed*, *hissed, wished* and *kissed*, and form an abstract *VERB-FRICATIVE+t* template by simply forgetting the beginnings of the individual words (which is what, after all, convey the different meanings). More importantly, this claim is at odds with the considerable evidence for the storage of such forms; i.e., frequency effects at the level of individual ready-inflected forms. Abstraction-as-forgetting is even more of a stretch at a sentence level, which would involve a speaker forgetting all of the individual words of a sentence – again at odds with findings of widespread lexical and n-gram effects – and retaining simply a highly abstract configuration of them.

Objection: But how do we “know” what to store? After all, we can’t know in advance that – for example – the colour of the speaker’s eyes or the shape of her face is linguistically irrelevant; we must be constrained to ignore such details and encode only linguistically-relevant ones.

Exemplar accounts assume that we do not know in advance what to store (e.g., Chandler, 2010). We take in as much information as we can from each episodic snapshot, albeit modified by attention (e.g., if our primary goal is to determine whether or not we recognize the speaker’s face, we might pay more attention to features in the visual than auditory modality). As language acquisition researchers, we generally assume that listeners store linguistic details – or even a narrow subset of “grammatically relevant” details – and disregard others. But if we were researchers of facial processing (a domain that has its own exemplar versus abstractions debate; e.g., Quiroga, 2017) we would focus instead on people’s ability to record eye colour and face shape, ignoring auditory information. And if we were researchers of body language, we would be interested in yet another set of features. Indeed, latent learning (as it is usually termed in animals), implicit learning and perceptual learning (in humans) are well-known phenomenon whereby participants show evidence of learning, and improved performance on subsequent tasks, even when they are viewing or processing stimuli with no particular goal – or even a goal that later turns out to be irrelevant – in mind (e.g., Thorpe, 1956; Reber, 1969, 1989; Kellman, 2002).

Even within the domain of language acquisition, what is grammatically irrelevant in one language (e.g., speaker gender, evidentiality, aspect, absolute position in geographical space) is critical in others. We cannot know what features will be important in advance, because the features that turn out to be relevant are determined retrospectively by some future goal; whether that is to produce a novel utterance of one’s own, recall the gist of what the speaker was saying, recognize her voice or face, or determine her attitude towards the other people present or the topic under discussion. At the linguistic level, it is exactly these “irrelevant” properties such as the precise sequence of words used (*n*-gram effects), the phonetic detail (speaker-effects), the particular meaning that the speaker seemed to have in mind at that instant (sidestepping the homophony problem when learning word meanings) that are key to the exemplar effects summarized throughout this paper.

Objection: Abstractions and exemplars are not the only two possibilities. Connectionist and naïve-discrimination-learning models undergo changes in response to each individual exemplar, but do not actually store them.

As Chandler (2010) notes, connectionism is an approach, not a single model. Many different types of connectionist models are possible. At one extreme, a connectionist model with a sufficiently large number of hidden units to represent each exemplar – particularly if supplemented with a mechanism for representing the order in which the exemplars were presented (e.g., Elman, 1990) – effectively *would* be an exemplar model. Each individual exemplar would be represented by a unique pattern of activation in the model's hidden layer. This is entirely compatible with exemplar accounts, which make no particular claims about the way in which exemplar storage is actually implemented. Indeed, unless one believes in Grandmother cells (e.g., Fodor & Pylyshyn, 1988), everyone agrees that the storage in the brain of any particular exemplar, concept etc. must be distributed across neurons in a way that is analogous to the distributed storage employed by connectionist (or "neural-net") models. At the other extreme, a connectionist model with very few hidden units would be forced, by design, to make abstractions that discard huge amounts of information, and so would fail to yield *n*-gram effects, speaker-recognition effects, and all of the other exemplar phenomena summarized in the present article. Connectionist models, then, are not an alternative to abstraction-based or exemplar models, but merely a computational framework in which either type of model (or, something in between) can be instantiated. The same can be said for naïve discrimination learning (NDL) models (e.g., Baayen, Hendrix & Ramscar, 2013) which are effectively simple connectionist models with no hidden layer: The cue-outcome pairings are either sufficiently rich as to represent each exemplar uniquely, in which case the NDL is just one particular instantiation of an exemplar model, or they are not, in which case the NDL cannot yield exemplar-level information that has been thrown away (e.g., phonetic information used for distinguishing different speakers).

Objection: Arguing that there is a meaningful difference between an exemplar account and a constructivist stored-abstractions account such as Tomasello (2003) is just splitting hairs: the only difference is whether the abstractions are stored or generated on the fly, and what difference does that really make to anything?

In fact, the accounts differ not only with regard to whether abstractions are stored or generated on the fly but, much more importantly, with regard to the nature of those abstractions. Constructivist stored-abstraction accounts posit very large, very abstract generalizations like the transitive [SUBJECT] [VERB] [OBJECT] construction, that can be used to produce novel utterances like *He kicked the ball* (see Section 4). An exemplar account differs in two important respects.

First, the emergent (rather than stored) generalization is at an extremely fine-grained level; more fine-grained than [SUBJECT] [VERB] [OBJECT], than [AGENT] [ACTION] [PATIENT], even than [KICKER] *kick* [KICKEE]. For example, depending on exactly what exemplar utterances have been stored, a speaker might generate *He kicked the ball* by analogy with *He kicked the tree*, *She kicked the stone*, and *He found the ball*.

Second, the generalization that emerges depends on exactly what the learner is trying to do at that moment. Suppose, for example, that, rather than trying to produce the utterance *He kicked the ball*, she is trying to figure out the meaning of an unknown word that she has just heard in the phrase *He kicked the widget*. The learner might recruit exactly the same exemplars – *He kicked the tree*, *She kicked the stone*, and *He found the ball* – but analogize across them in a very different way, placing most weight on semantic similarities between *tree, ball* and *ball*.

Objection: The assumption of representing language as stored episodic memory traces doesn't fit with either cognitive or neuroimaging findings regarding memory.
A traditional view is that memory consists of two largely separate systems, underpinned by different brain systems. For example, under Ullman's (2015) *declarative/procedural* model, the *declarative* (or explicit) system is characterized by conscious awareness (i.e., we can deliberately retrieve or "think about" these memories) and depends on the "the hippocampus and other medial temporal lobe (MTL)…and…neocortical regions, particularly…the temporal lobes". The *procedural* system (one of a number of implicit systems, alongside – for example – perceptual systems) is characterized by a lack of conscious awareness, and "involves a network of interconnected brain structures rooted in frontal/ basal-ganglia circuits, including frontal premotor and related regions, particularly BA 6 and BA 44". Declarative memory can be subdivided into *semantic* memory (e.g., the knowledge that Paris in the capital of France and *episodic* memory (e.g., the memory of a particular trip to Paris).

The exemplar account advocated here blurs these distinctions, because it places the language-learning burden squarely on the shoulders of *episodic* memory, despite the fact that we usually think of (native) linguistic knowledge as mainly *implicit*, though with some *semantic* memory for word meanings and, perhaps, morphosyntactic generalizations that are easily verbalized (e.g., "add -s to make a noun plural").

But this is not necessarily a problem. Although this *memory systems* framework is the one that is best known to researchers of child language acquisition (perhaps due to its popularization by Pinker & Ullman, 2002, in the context of the English past-tense debate), it is far from the only one. The *processing modes* framework (e.g., Roediger & McDermott's, 1993, *conceptual-perceptual* theory) argues that dissociations between performance on implicit and explicit memory tasks do not in fact reflect different memory systems, but merely different manners of processing. A third approach, the *component process* framework (e.g., Moscovitch, 1992) proposes that memory consists not of "a handful of memory systems or a couple of processing modes" but rather "dozens of different processing components... associated with different brain regions and…recruited in different combinations by memory" (e.g., Cabeza & Moscovitch, 2013: 50).

Child language researchers cannot be expected to play umpire in a debate in a different field, but the very existence of such a debate means that proposals in the domain of language acquisition should not be dismissed on the basis that they are at odds with a particular model of memory, particularly one that – at least according to some experts in this domain – is no longer the dominant model in the field.

Finally, at least one recent functional neuroimaging study (Mack et al, 2013; see also Nosofsky et al, 2012) has provided direct evidence for an exemplar- over prototype- account of categorization (albeit nonlinguistic categorization). Mack et al (2013) had participants perform a standard categorization task (involving different shapes of different colour, size and position) and used statistical models to generate, for each stimulus, for each participant, categorization predictions from a prototype and exemplar account. Although the behavioural predictions of the prototype and exemplar models were very similar, and could not differentiate between them, the internal states of the models were very different, and could be related to patterns of neural activation, as observed using fMRI. Mack et al found that, for 13/20 participants, the exemplar model yielded a better fit to the fMRI data than the prototype model, with only a single participant showing the opposite pattern (and six ties). More generally, Ashby and Rosedahl (2017) outline a proposal for how exemplar accounts could be instantiated neurobiologically. In summary, far from being biologically implausible, exemplar accounts exist as neurobiological models that provide a close fit to imaging data.

Objection: Saying that production/comprehension works "by analogy to stored exemplars" is just a big hand wave; there is no detail about how this analogy works. Certainly, an exemplar account of language acquisition (or, indeed, of any individual language acquisition phenomenon) will require details of the analogical process to be spelled out; and spelled out in sufficient detail that the account can be implemented and tested as a computational model. In domains outside of language acquisition, this has already been done. Indeed, as we have just seen, for nonlinguistic categorization tasks, exemplar accounts have been spelled out in sufficient detail to make predictions regarding patterns of voxel activation in fMRI studies (Mack et al, 2013). Somewhat closer to home, in the domain of Natural Language Processing, exemplar-based models (also known as case-, instance- or memory-based in this literature) have been formalized for tasks such as syllabification, grapheme-to-phoneme conversion and morphological analysis; see Daelemans & Van den Bosch, 2010, for a review).

For most domains of language acquisition, particularly core morphosyntax (Section 4), this process of mathematical formalization has not even begun[7]. However, for the very few, circumscribed domains for which a detailed exemplar account has been worked out (e.g., English past-tense acquisition), models based on the principle of analogy across stored exemplars have been very successful (see Chandler, 2010, for a review; see also Keuleers, Sandra, Daelemans Gillis, Durieux & Martens, 2007, on Dutch plurals, and Baayen & del Prado Martín, 2005, on semantic neighbourhoods in past-tense formation). My hope in writing this review is that colleagues will be inspired to begin exactly this kind of work, particularly in the domain of core morphosyntax.

In the meantime, it is important to be clear that stored abstraction accounts do not have any clear advantage here. Constructivist accounts such as that set out by Tomasello (2003) (or, indeed, by Ambridge & Lieven, 2015) also face the challenge of specifying the analogical processes that give rise to abstractions (in their case, stored abstractions). If anything, the challenge is greater, since such accounts need to explain the formation of very high-level abstractions (e.g., [SUBJECT] [VERB] [OBJECT], or, at least [AGENT] [ACTION] [PATIENT]) from input exemplars that have very little in common. Generativist accounts have an advantage in that they do not need to account for the formation of these high-level abstractions (e.g., [SUBJECT] [VERB] [OBJECT]), which have an important innate basis, and so are fully formed as soon as the relevant parameters have been set (e.g., Wexler, 1998). However, with this advantage comes a significant cost: If children are operating with such high-level abstractions from such an early age, it is very difficult to explain input effects, whether at the levels of *N*-grams, frequent sentence-level utterance patterns (e.g., *He's X-ing it*), the ability to identify different speakers and so on.

So, yes, there are no implemented exemplar models of how speakers produce particular utterances with the form [SUBJECT] [VERB] [OBJECT]; but there are no constructivist or generativist stored-abstraction ones either. The closest on the constructivist side is probably the dual-path model of Chang (2002; Chang, Dell & Bock, 2006); but, in the

---

[7] Although it is very far from a complete exemplar account of syntax acquisition, Schütze, Walsh, Möbius and Wade (2007) demonstrate that, for a toy SVO grammar with five nouns and five verbs, a simple exemplar model can distinguish unseen but grammatical strings (e.g., *I love coffee*) from unseen, ungrammatical strings (e.g., **I tea drink*), using pseudo-semantic representations consisting of the distributional context for each lexical item (e.g., *tea* and *coffee* tend to occur immediately to the right of *love, make* and *drink*, not *I*, *you* and *he*). Bod (2006) outlines a more ambitious exemplar-based model of syntax, albeit one that requires parsed input, and so works at a considerably higher level of abstraction that true exemplar models.

form of its semantic/event role representations, the model is already given at least one crucial piece of information that an exemplar account must deduce by analogy. On the generativist side, computational parameter-setting models such as that of Sakas and Fodor (2012) operate at a much higher level, investigating only how different possible orderings of categories such as [SUBJECT] [VERB] and [OBJECT] could be triggered; the categories themselves are already known.

That said, although only a handful of true exemplar models exist, other computational models that share some of their key assumption have achieved notable successes. We have already discussed connectionist models that instantiate exemplar storage and analogy, at least at a particular level of granularity. To these, we can add non-connectionist models such as MOSAIC (Freudenthal, Pine & Gobet, 2010), the chunk-based learner (e.g., McCauley & Christiansen, in press) and PUDDLE (Monaghan & Christiansen, 2010) that simulate acquisition phenomena in the domains of inflectional morphology (see Section 2), syntax (Section 4) and phonology (Section 5) on the basis of stored exemplars (though not, except in a very limited way, analogy across them).

Perhaps the best example, though, is Google Translate which, in its current incarnation (Johnson et al, 2016), works by storing surface exemplar strings – both words and phrases – each paired with a distributional vector representing its meaning, independent of the language in question (a crude approximation of the semantic "language of thought"). This allows the system to translate pairs that it has never seen (e.g., when trained on Japanese-English and Korean-English pairs, it can translate from Japanese to Korean). In many respects, Google Translate is an exemplar model of language production: When given a message to convey (i.e., a distributional meaning vector) it searches for stored exemplars with similar meanings, each paired with a particular surface string, and generates an output by analogizing across them. It is easy to find fault with Google Translate, particularly when it is thrown by phenomena that are the bread-and-butter of theoretical linguistics such as gender and number (example from Hofstadter, 2018):

> In their house, everything comes in pairs. There's his car and her car, his towels and her towels, and his library and hers. [English]
> Dans leur maison, tout vient en paires. Il y a sa voiture et sa voiture, ses serviettes et ses serviettes, sa bibliothèque et les siennes. [French]

But where are the alternatives; translation engines that work by parsing the input sentences in terms of stored abstractions – syntactic constituent structure or sentence-level constructions – then spelling them out using the equivalent structures in the target language? They don't exist, because they don't work. Researchers in machine translation have learned a lesson that should be heeded by researchers in child language acquisition: exemplar models are the only game in town.

References

Abbot-Smith, K., & Tomasello, M. (2006). Exemplar-learning and schematization in a usage-based account of syntactic acquisition. *The Linguistic Review*, *23*(3), 275-290.

Abbot-Smith, K., Lieven, E., & Tomasello, M. (2001). What preschool children do and do not do with ungrammatical word orders. *Cognitive Development, 16*(2), 679-692.

Abbot-Smith, K., & Behrens, H. (2006). How known constructions influence the acquisition of other constructions: The German passive and future constructions. *Cognitive Science*, *30*(6), 995-1026.

Abney, S. P. (1987). The English noun phrase in its sentential aspect (Doctoral dissertation, Massachusetts Institute of Technology).

Aguado-Orea, J., & Pine, J. M. (2015). Comparing different models of the development of verb inflection in early child Spanish. *PloS one*, *10*(3), e0119613.

Akhtar, N. (1999). Acquiring basic word order: evidence for data-driven learning of syntactic structure. *Journal of Child Language, 26*, 339-356.

Albright, A., & Hayes, B. (2003). Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition, 90*, 119-161.

Ambridge, B. (2010). Children's judgments of regular and irregular novel past tense forms: New data on the dual- versus single-route debate. *Developmental Psychology 46(6)*, 1497-1504.

Ambridge, B. & Lieven, E.V.M., (2011) *Child Language Acquisition: Contrasting theoretical approaches*. Cambridge, UK: Cambridge University Press.

Ambridge, B., & Lieven, E.V.M. (2015). A Constructivist account of child language acquisition. In B. MacWhinney and W. O'Grady (Eds). *Handbook of Language Emergence.* Hoboken, NJ: Wiley Blackwell. (pp. 478-510).

Ambridge, B., Bidgood, A., Pine, J.M., Rowland, C.F. & Freudenthal, D. (2016). Is passive syntax semantically constrained? Evidence from adult grammaticality judgment and comprehension studies. *Cognitive Science, 40*(6) 1435-1459

Ambridge, B., Bidgood, A., Twomey, E., Pine, J.M., Rowland, C.F. & Freudenthal, D. (2015). Preemption versus Entrenchment: Towards a construction-general solution to the problem of the retreat from verb argument structure overgeneralization. *PLoS ONE* 10(4): e0123723. doi:10.1371/journal.pone.0123723

Ambridge, B., Pine, J. M., Rowland, C. F., Jones, R.L, & Clark, V. (2009). A semantics-based approach to the 'no negative-evidence' problem *Cognitive Science*, 33(7), 1301-1316.

Ambridge, B., Pine, J.M., Rowland, C.F. & Young, C.R. (2008) The effect of verb semantic class and verb frequency (entrenchment) on children's and adults' graded judgements of argument-structure overgeneralization errors. *Cognition. 106(1), 87-129.*

Arnon I, Snider N (2010) More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language, 62*: 67–82. doi: 10.1016/j.jml.2009.09.005.

Arnon, I. & Clark, E. V. (2011). When 'on your feet' is better than 'feet': Children's word production is facilitated in familiar sentence-frames. *Language Learning and Development, 7*, 107-129.

Ashby, F. G., & Rosedahl, L. (2017). A neural interpretation of exemplar theory. *Psychological Review*, *124*(4), 472-482.

Aylett, M., & Turk, A. (2006). Language redundancy predicts syllable duration and the spectral characteristics of vocalic syllable nuclei. *Journal of the Acoustical Society of America, 119*, 3048–3058.

Baayen, R. H., Hendrix, P., & Ramscar, M. (2013). Sidestepping the combinatorial explosion: An explanation of n-gram frequency effects based on naive discriminative learning. *Language and Speech*, *56*(3), 329-347.

Baayen, R. H., & del Prado Martín, F. M. (2005). Semantic density and past-tense formation in three Germanic languages. *Language*, 666-698.

Bannard, C., & Matthews, D. (2008). Stored word sequences in language learning: The effect of familiarity on children's repetition of four-word combinations. *Psychological Science*, *19*(3), 241-248.

Bell, A., Brenier, J. M., Gregory, M., Girand, C., & Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language*, *60*(1), 92-111.

Blything, R.P., Ambridge, B., & Lieven, E.V.M. (2017). Children's acquisition of the English past-Tense: Evidence for a single-route account from novel verb production data. *Cognitive Science* DOI: 10.1111/cogs.12581.

Bock, K., & Loebell, H. (1990). Framing sentences. *Cognition, 35*(1), 1–39.

Bod, R. (2006). Exemplar-based syntax: How to get productivity from examples. *The Linguistic review*, *23*(3), 291-320.

Branigan, H. P., & Pickering, M. J. (2017a). An experimental approach to linguistic representation. *Behavioral and Brain Sciences, 40*.

Brooks, P. J., & Tomasello, M. (1999). How children constrain their argument structure constructions. *Language, 75*(4), 720-738.

Brown, R. (1973). *A First language: the early stages*. Cambridge, Mass.: Harvard University Press.

Bybee, J., & Scheibman, J. (1999). The effect of usage on degrees of constituency: the reduction of don't in English. *Linguistics, 37*(4), 575-596.

Cabeza, R., & Moscovitch, M. (2013). Memory systems, processing modes, and components: functional neuroimaging evidence. *Perspectives on Psychological Science*, *8*(1), 49-55.

Chandler, S. (2010). The English past tense: Analogy redux. *Cognitive Linguistics*, *21*(3), 371-417.

Chang, F., Bock, K., & Goldberg, A. E. (2003). Can thematic roles leave traces of their places?. *Cognition, 90*(1), 29-49.

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

Chang, F., Dell, G. S., Bock, K., & Griffin, Z. M. (2000). Structural priming as implicit learning: A comparison of models of sentence production. *Journal of Psycholinguistic Research*, *29*(2), 217-230.

Chater, N. & Christiansen, M.H. (in press). Language acquisition as skill learning. *Current Opinion in Behavioral Sciences*.

Childers, J. B., & Paik, J. H. (2009). Korean-and English-speaking children use cross-situational information to learn novel predicate terms. *Journal of Child Language*, *36*(1), 201-224.

Childers, J. B., & Tomasello, M. (2001). The role of pronouns in young children's acquisition of the English transitive construction. *Developmental Psychology, 37*(6), 739-748.

Chomsky, N. (1980). In Piatelli-Palmarini, M. (1980). *Language and Learning: The debate between Jean Piaget and Noam Chomsky*. Cambridge, MA: Harvard University Press.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. Cambridge, MA: MIT Press.

Conwell, E. (2017). Token Frequency Effects in Homophone Production: An Elicitation Study. *Language and Speech*, 0023830917737108.

Craik, F. I., & Kirsner, K. (1974). The effect of speaker's voice on word recognition. *The Quarterly Journal of Experimental Psychology*, *26*(2), 274-284.
Crain, S., & Nakayama, M. (1987). Structure dependence in grammar formation. *Language, 63*, 522-543.
Croft, W., & Cruse, A. (2004). *Cognitive Linguistics*. Cambridge, UK: Cambridge University Press.
Dąbrowska, E. (2018). Experience, aptitude and individual differences in native language ultimate attainment. *Cognition*, *178*, 222-235.
Dąbrowska, E., & Szczerbinski, M. (2006). Polish children's productivity with case marking: the role of regularity, type frequency, and phonological diversity. *Journal of Child Language, 33*(3), 559-597.
Dąbrowska, Ewa (2008) "The effects of frequency and neighbourhood density on adult native speakers' productivity with Polish case inflections: An empirical test of usage-based approaches to morphology". *Journal of Memory and Language* 58, 931-951.
Daelemans, W. and Van Den Bosch, A. (2010). Memory-Based Learning. In *Handbook of computational linguistics and natural language processing* (eds A. Clark, C. Fox and S. Lappin). doi:10.1002/9781444324044.ch6.
Dittmar, M., Abbot-Smith, K., Lieven, E., & Tomasello, M. (2008). Young German children's early syntactic competence: A preferential looking study. *Developmental Science*, *11*(4), 575-582.
Dodson, K., & Tomasello, M. (1998). Acquiring the transitive construction in English: The role of animacy and pronouns. *Journal of Child Language, 25*, 555-574.
Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*(2), 179-211.
Engelmann, F., Granlund, S., Kolak, J., Zreder, M., Ambridge, B., Pine, J.M., Theakston, A.L., & Lieven, E.V.M. (submitted). How the input shapes the acquisition of verb morphology: Elicited production and computational modelling in two highly inflected languages.
Erk, K. & Padó, S. (2010). Exemplar-based models for word meaning in context. *Proceedings of the Association for Computational Linguistics 2010 Conference Short Papers*, (pp. 92–97).
Fodor, J., & Lepore, E. (1996). The red herring and the pet fish: Why concepts still can't be prototypes. *Cognition*, *58*(2), 253-270.
Fodor, J. A., and Pylyshyn, Z. (1988). Connectionism and cognitive architecture: a critical analysis. *Cognition* 28, 3–71.
Freudenthal, D., Pine, J., & Gobet, F. (2010). Explaining quantitative variation in the rate of Optional Infinitive errors across languages: a comparison of MOSAIC and the Variational Learning Model. *Journal of child language*, *37*(3), 643-669
Gertner, Y., Fisher, C., & Eisengart, J. (2006). Learning words and rules: Abstract knowledge of word order in early sentence comprehension. *Psychological Science, 17*(8), 684-691.
Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language.* New York: Oxford University Press.
Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, *105*(2), 251.
Goldinger, S.D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1166-1183.
Goldwater, M. B., Tomlinson, M. T., Echols, C. H., & Love, B. C. (2011). Structural priming as structure-mapping: children use analogies from previous utterances to guide sentence production. *Cognitive Science*, *35*(1), 156-170.

Gordon, K. R., McGregor, K. K., Waldier, B., Curran, M. K., Gomez, R. L., & Samuelson, L. K. (2016). Preschool children's memory for word forms remains stable over several days, but gradually decreases after 6 months. *Frontiers in Psychology,* https://doi.org/10.3389/fpsyg.2016.01439

Guenther, F. H., & Gjaja, M. N. (1996). The perceptual magnet effect as an emergent property of neural map formation. *The Journal of the Acoustical Society of America*, *100*(2), 1111-1121.

Hartshorne, J. K., & Ullman, M. (2006). Why girls say "holded" more than boys. *Developmental Science, 9*(1), 21-32.

Hay, J. (2018). Sociophonetics: The role of words, the role of context, and the role of words in context. *Topics in Cognitive Science.*

Hay, J., & Bresnan, J. (2006). Spoken syntax: The phonetics of giving a hand in New Zealand English. *The Linguistic Review*, *23*(3), 321-349.

Hofstadter (2018). https://www.theatlantic.com/technology/archive/2018/01/the-shallowness-of-google-translate/551570/

Houston, D. M. & Jusczyk, P. W. (2000). The role of talker-specific information in word segmentation by infants. Journal of Experimental Psychology: Human Perception and Performance 26, 1570–82.

Houston, D. M. & Jusczyk, P. W. (2003). Infants' long-term memory for the sound patterns of words and voices. Journal of Experimental Psychology: Human Perception and Performance 29, 1143–54

Janssen, N. & Barber, H.A. (2012) Phrase frequency effects in language production. PLoS ONE 7(3): e33202. doi:10.1371/journal.pone.0033202.

Johnson, K. (2007). Decisions and mechanisms in exemplar-based phonology. In Sole, M., Speeter Beddor, P., & Ohala, M. *Experimental approaches to phonology*. Oxford: Oxford University Press. (pp.25-40).

Johnson, M., Schuster, M., Le, Q. V., Krikun, M., Wu, Y., Chen, Z., ... & Hughes, M. (2016). Google's multilingual neural machine translation system: enabling zero-shot translation. arXiv preprint arXiv:1611.04558.

Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. in Bybee, Joan and Paul Hopper (eds.). 2000. Frequency and the emergence of linguistic structure. Amsterdam: John Benjamins (pp. 229-254).

Jusczyk, P. W. (1993). From general to language-specific capacities: The WRAPSA model of how speech perception develops. *Journal of Phonetics, 21*(1-2), 3-28.

Kellman, P. J. (2002). Perceptual learning. *Stevens' Handbook of Experimental Psychology*.

Keuleers, E., D. Sandra, W. Daelemans, S. Gillis, G. Durieux, & E. Martens (2007), Dutch plural inflection: The exception that proves the analogy, *Cognitive Psychology 54*(4):283–318.

Kidd, E., & Lum, J. A. (2008). Sex differences in past tense overregularization. *Developmental Science*, *11*(6), 882-889.

Kim, S.N. & T. Baldwin (2006), Interpreting semantic relations in noun compoundsvia verb semantics, in Proceedings of the COLING/ACL 2006 Main Conference Poster Sessions, Sydney, Australia, (pp. 491–498).

Kirjavainen, M., Nikolaev, A., & Kidd, E. (2012). The effect of frequency and phonological neighbourhood density on the acquisition of past tense verbs by Finnish children. *Cognitive Linguistics*, *23*(2), 273-315.

Kjærbæk, L., dePont Christensen, R., & Basbøll, H. (2014). Sound structure and input frequency impact on noun plural acquisition: Hypotheses tested on Danish children across different data types. *Nordic Journal of Linguistics*, *37*(1), 47-86.

Krajewski, G., Theakston, A. L., Lieven, E. V., & Tomasello, M. (2011). How Polish children switch from one case to another when using novel nouns: Challenges for models of inflectional morphology. *Language and Cognitive Processes*, *26*(4-6), 830-861.

Krug, M. (1998). String frequency. A cognitive motivating factor in coalescence, language processing, and linguistic change. *Journal of English Linguistics, 26*, 286–320.

Kuhl, P. K. (1991). Human adults and human infants show a “perceptual magnet effect” for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics*, *50*(2), 93-107.

Kunnari, S., Savinainen-Makkonen, T., Leonard, L. B., Mäkinen, L., Tolonen, A. K., Luotonen, M., & Leinonen, E. (2011). Children with specific language impairment in Finnish: The use of tense and agreement inflections. *Journal of Child Language*, *38*(5), 999-1027.

Labov, W. (1963). The social motivation of a sound change. *Word*, *19*(3), 273-309.

Langacker, R.W. (2009). The emergence of structure from usage. Unpublished manuscript, University of California, San Diego.

Leonard, L. B., Caselli, M. C., & Devescovi, A. (2002). Italian Children's Use of Verb and Noun Morphology during the Preschool Years. *First Language, 3*(66), 287-304.

Lieberman, P. (1963). Some effects of semantic and grammatical context on the production and perception of speech. *Language and Speech, 6*(3), 172-187.

Lieven, E., Salomo, D., & Tomasello, M. (2009). Two-year-old children's production of multiword utterances: A usage-based analysis. *Cognitive Linguistics*, *20*(3), 481-507.

Loewenstein, J., & Gentner, D. (2001). Spatial mapping in preschoolers: Close comparisons facilitate far mappings. *Journal of Cognition and Development*, *2*(2), 189-219.

Love, B.C. (2013). Categorization. In K.N. Ochsner and S.M. Kosslyn (Eds.) *Oxford Handbook of Cognitive Neuroscience*, (pp.342-358). Oxford University Press.

Mack, M. L., Preston, A. R., & Love, B. C. (2013). Decoding the brain’s algorithm for categorization from its neural implementation. *Current Biology*, *23*(20), 2023-2027.

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.

Maratsos, M. (2000). More Overregularizations after All: New Data and Discussion on Marcus, Pinker, Ullman, Hollander, Rosen & Xu. *Journal of Child Language, 27*(1), 183-212.

Marchman, V. A. (1997). Children's productivity in the English Past Tense: The Role of Frequency, Phonology and Neighborhood Structure. *Cognitive Science, 21*(3), 283-304.

Marchman, V. A., & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language*, *21*(2), 339-366.

Marchman, V. A., Wulfeck, B., & Weismer, S. E. (1999). Morphological productivity in children with normal language and SLI: A study of the English Past Tense. *Journal of Speech, Language and Hearing Research, 42*, 206-219.

Maslen, R. J. C., Theakston, A. L., Lieven, E. V. M., & Tomasello, M. (2004). A dense corpus study of past tense and plural overregularization in English. *Journal of Speech, Language and Hearing Research, 47*, 1319-1333.

Matthews, D., & Bannard, C. (2010). Children’s Production of Unfamiliar Word Sequences Is Predicted by Positional Variability and Latent Classes in a Large Sample of Child-Directed Speech. *Cognitive science*, *34*(3), 465-488.

Matthews, D., Lieven, E., Theakston, A. L., & Tomasello, M. (2004). The role of frequency in the acquisition of English word order. *Cognitive Development, 20*, 121-136.

Matthews, D., Lieven, E., Theakston, A., & Tomasello, M. (2007). French children's use and correction of weird word orders: A constructivist account. *Journal of Child Language, 34*(2), 381-409.

Mattys, S. L., & Jusczyk, P. W. (2001). Phonotactic cues for segmentation of fluent speech by infants. *Cognition*, *78*(2), 91-121.

McClure, K., Pine, J. M., & Lieven, E. V. M. (2006). Investigating the abstractness of children's early knowledge of argument structure. *Journal of Child Language, 33*, 693–720.

McDonald, S. A., & Shillcock, R. C. (2003). Eye movements reveal the on-line computation of lexical probabilities. *Psychological Science, 14*, 648–652.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, *85*(3), 207.

Monaghan, P., & Christiansen, M. H. (2010). Words in puddles of sound: Modelling psycholinguistic effects in speech segmentation. *Journal of child language*, *37*(3), 545-564.

Morgan, J. L., Singh, L., Bortfeld, H., Rathbun, K., & White, K. (2001). Effects of speech and sentence position on infant word recognition. Paper presented at the Boston University Conference on Language Development, Boston, MA.

Moscovitch, M. (1992). Memory and working-with-memory: A component process model based on modules and central systems. *Journal of Cognitive Neuroscience, 4*, 257–267.

Nastase, V., J. Sayyad-Shiarabad, M. Sokolova, & S. Szpakowicz (2006), Learning noun-modifier semantic relations with corpus-based and wordnet-based features, in *Proceedings of the Twenty-First National Conference on Artificial Intelligence and the Eighteenth Innovative Applications of Artificial Intelligence Conference*, AAAI Press (pp.781-786).

Newmeyer, F. J. (2003). Grammar is grammar and usage is usage. *Language*, *79*(4), 682-707.

Nosofsky, R. M., Little, D. R., & James, T. W. (2012). Activation in the neural network responsible for categorization and recognition reflects parameter changes. *Proceedings of the National Academy of Sciences*, *109*(1), 333-338.

Nosofsky, R. M. (2016). An exemplar-retrieval model of short-term memory search: Linking categorization and probe recognition. In *Psychology of Learning and Motivation* (Vol. 65, pp. 47-84). New York: Academic Press.

Oakes, L. M., Coppage, D. J., & Dingel, A. (1997). By land or by sea: the role of perceptual similarity in infants' categorization of animals. *Developmental Psychology*, *33*(3), 396.

Palmeri, T. J., Goldinger, S. D., & Pisoni, D. B. (1993). Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(2), 309.

Perry, L. K., Samuelson, L. K., Malloy, L. M., & Schiffer, R. N. (2010). Learn locally, think globally: Exemplar variability supports higher-order generalization and word learning. *Psychological Science*, *21*(12), 1894-1902.

Pinker, S. (1987). The bootstrapping problem in language acquisition. In B. MacWhinney (ed.), Mechanisms of language acquisition (pp. 399-442). Hillsdale, NJ: Erlbaum.

Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT.

Pinker, S., & Ullman, M. T. (2002). The past and future of the past tense. *Trends in cognitive sciences*, *6*(11), 456-463.

Pluymaekers, M., Ernestus, M., & Baayen, R. (2005). Articulatory planning is continuous and sensitive to informational redundancy. *Phonetica*, *62*(2-4), 146-159.

Posner, M. I., & Keele, S. W. (1970). Retention of abstract ideas. *Journal of Experimental Psychology*, *83*(2, Pt.1), 304-308.

Prasada, S., & Pinker, S. (1993). Generalisation of regular and irregular morphological patterns. *Language and Cognitive Processes, 8*(1), 1-56.

Prince, A., & Smolensky, P. (2008). *Optimality Theory: Constraint interaction in generative grammar.* John Wiley & Sons.

Quiroga, R. Q. (2017). How do we recognize a face?. *Cell*, *169*(6), 975-977.

Ramscar, M. & Port, R (2015) Categorization (without categories). In E. Dawbroska & D. Divjak (Eds.), H*andbook of Cognitive Linguistics,* The Hague: De Gruyter Mouton.

Ramscar, M., & Port, R. F. (2016). How spoken languages work in the absence of an inventory of discrete units. *Language Sciences*, *53*, 58-74.

Rasanen, S. H. M., Ambridge, B. & Pine, J. M. (2016). An elicited-production study of inflectional verb morphology in child Finnish. *Cognitive Science, 40*(7), 1704- 1738.

Reber, A. S. (1969). Transfer of syntactic structure in synthetic languages. *Journal of Experimental Psychology*, *81*(1), 115.

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, *118*(3), 219.

Roediger, H. L., & McDermott, K. B. (1993). Implicit memory in normal human subjects. In F. Boller & J. Grafman (Eds.), Handbook of neuropsychology (Vol. 8, pp. 63–131). Amsterdam: Elsevier.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*(4), 573-605.

Rowland, C.F., Chang, F., Ambridge, B., Pine, J. M., & Lieven, E.V.M. (2012). The development of abstract syntax: Evidence from structural priming and the lexical boost. *Cognition, 125*(1) 49–63.

Rubino, R. B., & Pine, J. M. (1998). Subject-verb agreement in Brazilian Portuguese: what low error rates hide. *Journal of Child Language, 25*(1), 35-59.

Saffran, J.R., Aslin, R.N., & Newport, E.L. (1996). Statistical learning by 8-month-olds. *Science*, 274, 1926-1928.

Sakas, W. G., & Fodor, J. D. (2012). Disambiguating syntactic triggers. *Language Acquisition, 19*(2), 83-143.

Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2003). Testing the abstractness of children's linguistic representations: lexical and structural priming of syntactic constructions in young children. *Developmental Science, 6*(5), 557-567.

Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2006). Structural priming as implicit learning in language acquisition: The persistence of lexical and structural priming in 4-year-olds. *Language Learning and Development, 2,* 27-50.

Savičiūtė, E., & Ambridge, B. (2018). The roles of word-form frequency and phonological neighbourhood density in the acquisition of Lithuanian noun morphology. *Journal of Child Language*, *45*(3), 641-672.

Schütze, H., Walsh, M., Möbius, B. & Wade. T. (2007): Towards a Unified Exemplar Theoretic Model of Phonetic and Syntactic Phenomena". Proceedings of the 29th Meeting of the Cognitive Science Society (CogSci 2007, Nashville, TN), pp. 1461-1466.

Siyanova-Chanturia, A., Conklin, K., & van Heuven, W.J.B. (2011) Seeing a phrase "time and again" matters: The role of phrasal frequency in the processing of multiword sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition 37*: 776.

Smith, D. J., & Minda, J. P. (2000). Thirty categorization results in search of a model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(1), 3-27.

Smoke, K. (1932) An objective study of concept formation. *Psychological Monographs XLII*(191): 1−46.
Sosa, A. V., & Stoel-Gammon, C. (2012). Lexical and phonological effects in early word production. *Journal of Speech, Language, and Hearing Research*, *55*(2), 596-608.
Sosa, A.V. & MacFarlane, J. (2002) Evidence for frequency-based constituents in the mental lexicon: Collocations involving the word of. *Brain and Language 83*: 227–236.
Standing, L. (1973). Learning 10000 pictures. Quarterly Journal of Experimental Psychology, 25(2), 207-222.
Stemberger, J. (1993). Vowel dominance in over-regularizations. *Journal of Child Language*, *20*(3), 503–521.
Stoel-Gammon, C. (2011). Relationships between lexical and phonological development in young children. *Journal of Child Language*, *38*(1), 1-34.
Tatsumi, T., Ambridge, B., & Pine, J. M. (2018). Disentangling effects of input frequency and morphophonological complexity on children's acquisition of verb inflection: An elicited production study of Japanese. *Cognitive Science*, *42*, 555-577.
Thorpe, W. H. (1956). *Learning and instinct in animals.* Cambridge, MA, US: Harvard University Press.
Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition, 74*(3), 209-253.
Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
Tremblay, A. and Baayen, R. H. (2010) Holistic processing of regular four-word sequences: A behavioral and ERP study of the effects of structure, frequency, and probability on immediate free recall. In D. Wood, *Perspectives on formulaic language: Acquisition and communication*. London: The Continuum International Publishing Group
Ullman, M. T. (2015). The declarative/procedural model: a neurobiological model of language learning, knowledge, and use. In Neurobiology of language (pp. 953-968).
Valian, V. (2014). Arguing about innateness. *Journal of Child Language*, *41*(S1), 78-92.
Wexler, K. (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua, 106*(1-4), 23-79.
Zaki, S. R., Nosofsky, R. M., Stanton, R. D., & Cohen, A. L. (2003). Prototype and exemplar accounts of category learning and attentional allocation: a reassessment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(6), 1160–1173.
Ziegler, J., & Snedeker, J. (2018). How broad are thematic roles? Evidence from structural priming. *Cognition*, *179*, 221-240.
Ziegler, J., Goldberg, A., & Snedeker, J. (2018). Passive priming requires function word overlap. Poster presented at the 31st Annual Meeting of the CUNY Conference on Human Sentence Processing, Davis, CA.
Ziegler, J., Snedeker, J., & Wittenberg, E. (2017). Priming is swell, but it's far from simple. *Behavioral and Brain Sciences, 40*, 44-45.