(De) Constructing Sentences

Ewa Dąbrowska
(Northumbria University, UK)

Abstract: Lexically specific units, i. e. , formulaic frames (*Where's Daddy? What's Mummy doing?*) and frames with slots (*e. g. Where's NP? Do you want to VP?*) are known to play a central role in language acquisition. Could such mechanisms also account for adult production? I argue that the types of representations and production mechanisms that constructivist language researchers have postulated for children can also explain the *basic mode* of adult language use, i. e. informal conversation. Viewing adult language in such terms not only captures the continuity between child and adult usage, but also helps to explain how speakers are able to produce and understand language as fast as they do: using preconstructed chunks saves processing effort. Arguably, adults also have more abstract linguistic representations. These, however, are acquired relatively late in acquisition, largely as a result of experience with written texts, and may not be acquired by all speakers.
Key words: lexically specific units; formulaic language; usage-based models; spoken language; language production

Any cognitively realistic theory of language must account for at least three basic facts: the fact that we are able to use language productively and produce novel sentences like "Familiar old clichés live virtually"; the fact that children are able to learn language from the input available to them; and the fact that we are able to process language in real time. We are able to process language extremely efficiently. In ordinary conversation, we produce something of the order of 150 words per minute, and the gaps between turns are on average about 500 milliseconds. If you think about it, this is the amount of time that the second interlocutor has to process the preceding utterance, interpret it, decide on a relevant response, formulate it, and implement the motor program to produce it. And it doesn't quite happen like that. We know that speech is processed as soon as it's heard. So we actually do a lot of processing while still listening to the utterance. And in fact, quite often speakers will respond to the previous interlocutor's point even before the first speaker has finished speaking, before she has finished her turn. So we do this extremely fast.

What makes the task even more difficult is that the speech stimulus is often degraded. This is illustrated by a study by Pollack and Pickett (1964), who recorded people in conversation and then spliced out single words and played them to people who didn't participate in the original conversation. They found that only about 50 percent of the words could be identified out of context. Furthermore, normal conversation often contains sentence fragments, false starts, and some sentences that are simply ungrammatical. Yet listeners are not fazed by this.

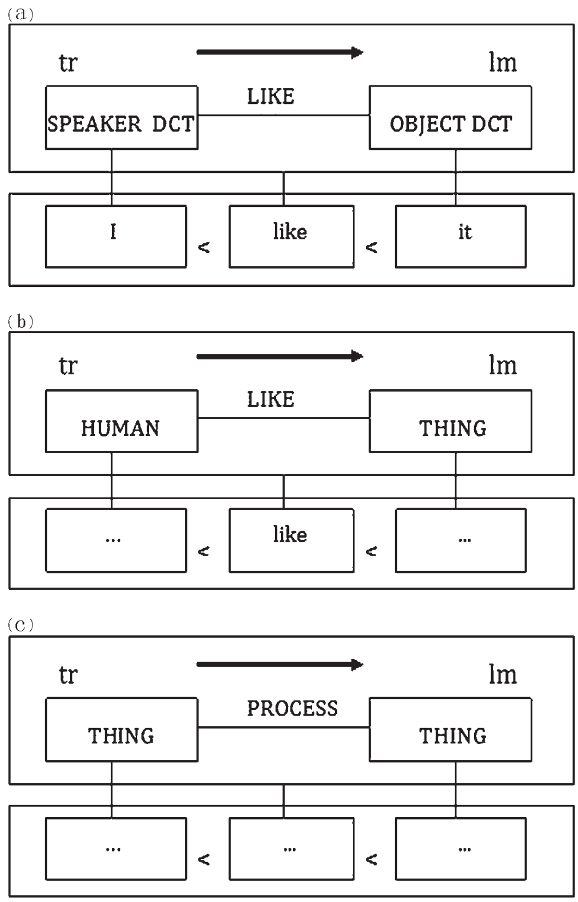
Generative linguistics has concentrated on productivity, and so it meets the first criterion.

However, Chomsky explicitly acknowledged that a generative grammar is unlearnable from the input (hence the need for Universal Grammar). And it doesn't do very well on processing. Most generative linguists are not interested in processing because it is regarded as an aspect of performance rather than competence; and the models that have been developed in that framework are not terribly useful when you are interested in language processing because they don't really tell you how you get from the sound to the meaning or vice versa and they are also computationally quite intractable.

Cognitive linguistics, on the other hand, does very well on all three requirements. We can explain how speakers are able to produce novel sentences by appealing to constructional schemas, and we can explain how children are able to learn them from the input available to them (see, for example Tomasello 2003, 2006, Dąbrowska and Lieven 2005). There is less cognitively-based work on processing: it is an area that has been neglected in cognitive linguistics. We do not really talk to psycholinguists. And we don't reflect very much on what people actually do in conversation. So I'll have a few suggestions to make in this area.

I'd like to start with some basic assumptions of the usage-based cognitive approaches. Linguistic knowledge consists of a network of constructions or form-meaning pairings. These can be simple or complex and they can be concrete or abstract. Crucially, abstract constructions have the same structure as the more specific units. Let's take a phrase like *I like it*. This phrase is quite frequent, and it's not unreasonable to suppose that it is available to many speakers as a preconstructed unit. It designates the relationship between the speaker and some object that they like. And you can represent it by three phonological units, *I*, *like*, and *it* which are arranged in a certain order (see Figure 1a). You can also have a somewhat more abstract unit which we may call a "low level schema" or "verb island" for the verb *like*; child language researchers often call such units "formulaic frames" or "frames with slots". This unit has the same overall structure, but it's specified in less detail so that the "liker" and the object liked are semantically more general, and there are slots at the phonological pole into which novel material can be inserted (Figure 1b). Finally, we have a fully general transitive construction, which is even more abstract in that all three slots are specified in very general terms semantically, and there are only slots at the phonological pole (Figure 1c). Crucially, all three constructions have the same overall structure.

Figure 1: A formulaic phrase (a), frames with slots (b), and an abstract schema (c)



To acquire a complex semantic unit such as the formulaic frame or the transitive schema, a learner must be able to do three things. First, he or she must be able to represent phonological and semantic properties of utterances. Secondly, the learner must be able to segment utterances and match chunks of phonology with chunks of semantic structure: they must be able to work out that in a sentence like *I like it*, the phonological form *I* corresponds to the speaker, *it* corresponds to the object liked, and *like* corresponds to the relationship between them. And thirdly, the learner must be able to form slots by generalizing over fillers attested in a particular frame. Crucially, all of this is necessary for lexical learning. So in order to learn a verb like *like*, you have to be able to form these sorts of abstractions. You can think of a low-level schema or frame with slots as a mini-grammar for the verb *like*, one that allows you to construct novel sentences with this verb. You can also think of it as the lexical representation of the verb. Crucially, the mechanisms necessary to acquire complex partially schematic units are the same as those needed to acquire relational words such as verbs (which, as I pointed out earlier, are effectively frames with slots). Children must be able to learn words because there is so much linguistic variation in the lexicon. And to explain how they are able to learn grammar, we don't need to assume any cognitive abilities over and above those that we know are necessary for lexical learning: we know that children have these.

I will now briefly discuss some earlier work I did jointly with Elena Lieven (Dąbrowska and Lieven 2005), and then move on to some new work on adult production. Dąbrowska and Lieven 2005 is an attempt to explain children's linguistic productivity in terms of two types of units, fixed phrases and frames with slots, and two operations, superimposition and juxtaposition.

The study used 8 corpora for 4 children who were recorded for 30 hours at the age of 2 and for 30 hours at the age of 3. We divided each of the corpora into a main corpus and a test corpus. The test corpus was the last transcript of the corpus, or, if the child didn't say very much in the last recording, the last two transcripts in the corpus. The main corpus comprised the remaining transcripts.

The main research question posed in the paper was whether we could derive the utterances in the test corpus from lexically specific units attested in the main corpus — in other words, whether we could "trace them back" to earlier utterances. The purpose of this was to develop a "child-friendly grammar", that is to say, a grammar that contains only lexically specific units (which are learnable from the input).

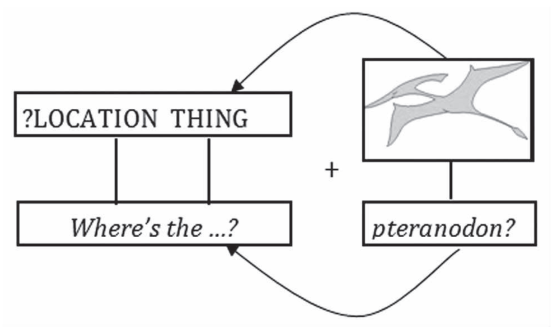
We identified two kinds of lexically specific units in the main corpus: fixed phrases and frames with slots. **Fixed phrases** are form-meaning pairings consisting of more than one word, such as *to the shop*. We assumed that recurrent form-meaning combinations were available to learners as units if they were attested in the main corpus at least twice. This may seem like a very unrealistic criterion: after all, it's unlikely that a learner will remember a phrase that they have only heard twice. I will return to this issue later: for now, bear in mind that we only have a partial record of children's linguistic experience — about 1% of the two-year-olds' linguistic experience and a mere 0.5% of the three-year-olds' experience.

Frames with slots such as *shall we … then?* (e.g. *shall we do some drawing then? shall we go back then?*) are also recurrent form-meaning combinations. In this case, the meaning is a suggestion that the speaker and the interlocutor do something together; and the form is *shall we …*

then? with a slot in the middle where you can specify the type of activity that you'd like to do together with your interlocutor. Again we assumed that these are available as units if they were attested at least twice in the main corpus with different fillers.

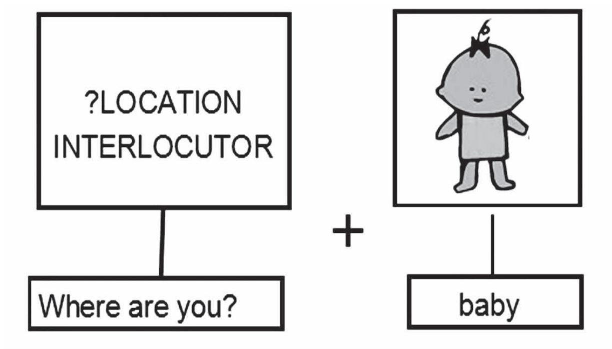
These units — fixed phrases and frames with slots — can be combined using two operations: superimposition or juxtaposition. **Superimposition** involves superimposing — or "merging" or "unifying" — two expressions so that the corresponding parts overlap. So there has to be some sort of correspondence between the semantic properties of the filler and the semantic properties of the slot into which it goes. Take the utterance *Where is the pteranodon*? produced by one of our three-year-olds. This can be assembled by superimposing two units, as shown in Figure 2. At the semantic pole, PTERANODON goes into the THING slot in the frame; at the same time, the phonological form *pteranodon* is inserted after *where's the*....

Figure 2: An example of superimposition



Juxtaposition involves putting two units together side by side where there is no direct syntactic relationship between them (cf. Figure 3). The relationship is basically paratactic, and the units can be combined in either order: for instance, when the units *where are you*? and *baby* are juxtaposed, the result can be either *where are you baby*? or *baby where are you*? Juxtaposition is quite restricted in English: basically you can only do this with vocatives and with some adverbials.

Figure 3: An example of juxtaposition



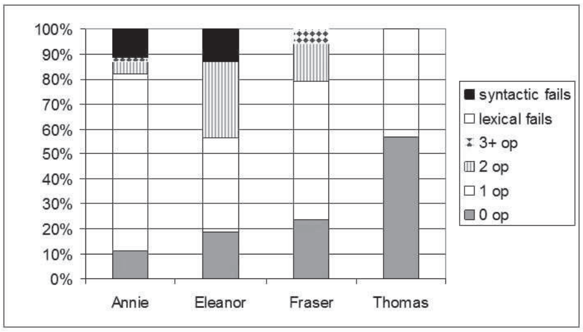
Let us now look at a few examples of how these two operations could be used to construct normal utterances. One of our three-year-olds produced the utterance *Shall we get them ready then*? We can derive this utterance by combining three chunks attested in the main corpus. *Get them AP* and *get NP ready* are superimposed such that the corresponding parts overlap, so *get* corresponds to *get*, *them* elaborates the NP slot in the second schema and *ready* elaborates the AP slot in the first schema. Notice that it is also possible to derive the expression *get them ready* by superimposing the adjective *ready* over the AP slot in the first schema, or by superimposing the pronoun *them* over the NP slot in the second schema. In this analysis, we assumed that language learners will always use the largest units available. That's quite a controversial assumption, and there are two motivations for it. One is that it might be easier for learners to combine units if they overlap. This is best explained by using an analogy. Imagine that you are assembling a jigsaw puzzle in which each piece has a picture of part of the adjoining piece attached to it. This makes it easier to work out how to put the two pieces together. The other reason is that if you use smaller units, you do get occasional errors of commission: for instance, if you substitute *they* instead of *them* in the NP slot in *get NP ready*, you get *get they ready*, which of course is ungrammatical. Children sometimes make errors like this. However, they can avoid making such errors if they use larger units.

Once you've got *get them ready*, this expression can go into the VP slot in the third schema (*shall we VP then?*) and you get *shall we get them ready then?*

Let us look at a second example, the utterance *you don't need to go to the bathroom, do you?*. We've got three chunks here: *go to the bathroom*, *don't need to VP*, and *you don't VP, do you?*. You can superimpose these three units and get *You don't need to go to the bathroom, do you?*. Again, there are alternative ways of doing it, so you might just use *need to VP* rather than *don't need to VP*, or you might use a shorter chunk instead of *go to the bathroom*, such as *go to the NOUN* or *go to NP* or *go LOCATION*. I'll come back to this point later.

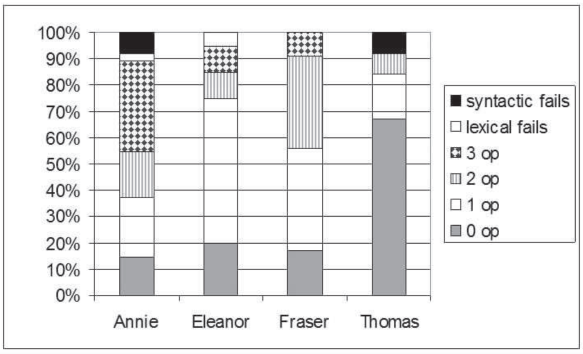
Figure 4 shows the results of the analysis for two-year-olds. As we can see, the majority of the children's utterances are either what we call zero operation utterances or single operation utterances. **Zero operation utterances** are utterances which are identical to lexically specific units which occurred at least twice in the main corpus, and therefore, by hypothesis, they are available as preconstructed units. **Single operation utterances** involve just one operation. Together these two types account for the majority of the children's utterances. There are also a few two- and three-operation utterances in some of the children. And finally, there are some "syntactic fails," i. e. cases where we cannot derive the child's utterances from the units in the main corpus.

Figure 4: Traceback results for 2-year-olds (based on Dąbrowska and Lieven 2005)



For three year olds, the results are quite similar (see Figure 5). The only difference is that we get more two-, three-, and, in Annie's case, also four-operation utterances. Overall, between 87% and 100% of the children's utterances can be derived using the lexically specific chunks in the main corpus.

Figure 5: Traceback results for 3-year-olds (based on Dąbrowska and Lieven 2005)



A reasonable question to ask at this stage is "What about the fails?" In other words, what about the utterances that cannot be derived using lexically specific chunks, which account for between 0 and 13% of the utterances in the test corpus? There are a number of reasons why such utterances appear in the corpus. First, they may be simply speech errors. Secondly, they may be instances of the child trying to say something that he or she has not yet learned how to say, so the child might be innovating. There is some evidence that at least some of the underivable utterances are either speech errors or instances of innovation: 62% of the fails are ungrammatical, compared to just 20% of the successful derivations.

Another possibility is that they are sampling artifacts. Our corpus comprises approximately 1% of the two-year-olds' linguistic experience and approximately 0.5% of the three-year-olds' linguistic experience. It is possible that we would be able to derive more of the target utterances if we had a larger corpus of the child's prior linguistic experience. Notice that the same case can be made for lexical fails. Lexical fails are utterances which could not be derived because a particular word did not occur at least two times in the main corpus. Obviously if a child used a word which is conventional in the adult community, he or she must have heard it before. The same logic applies to larger units.

Finally, the fails may be the result of application of more abstract rules. This is unlikely to be the case, however, for two reasons. First, the fails are not more complex than the successful derivations; and secondly, most of them are what we called 'near misses': that is to say, you could derive them if you relaxed the criteria just a little bit (see Dąbrowska and Lieven 2005).

So we can provisionally conclude that children's creativity involves recycling memorized chunks, or what Mike Tomasello calls "cut-and-paste". Crucially, I'd like to stress that in order to do cut-and-paste, you must have some understanding of the role of the component units in the schema. So you must have some kind of grammatical knowledge, it's just a different kind of grammatical knowledge from what has normally been assumed in the past.

Let us now consider adult production. Could the method described earlier possibly work for adults as well? Is it realistic to assume that adults also use "cut-and-paste"? There is a substantial amount of research suggesting that multiword chunks and holistic processing are also important for adults. It comes from two main traditions. One is work on phraseology and idioms; the other is work on processing shortcuts, which attempts to explain how people are able to process language as quickly as they are. But most of the work in these traditions considers holistic processing as a special mode of production which complements the normal or analytic route. As I've already pointed out, there is also a considerable amount of work on the use of formulaic phrases in language acquisition (both first and second language acquisition). However, most of this research sees holistic processing as a developmental phenomenon, something that learners eventually progress beyond.

There is quite a lot of evidence that even adults are quite good at remembering phrases and sentences which they have heard before. I will just briefly mention one study, by Gurevich et al. (2010). This paper describes a series of very carefully designed experiments in which participants listened to a story, and then they were given either a recall test or a recognition test. There were two versions of the story which had more or less the same content but differed in form. The sentences from the other version of the story were used as foils in the recognition task, and as controls in the recall task, to estimate approximately how often speakers accidentally produce the target utterance when retelling the story. The results showed that participants were able to recognize sentences that they have heard only once at above chance level — between 60 and 73 percent of the time, depending on the experiment. In the recall task, they reproduced between 11 and 22 percent of the sentences verbatim. They were even able to reproduce some of the sentences verbatim when tested a week later.

There is also evidence that speakers find frequent multiword combinations easier to process than less frequent ones. For instance, frequent word combinations are read faster (Tremblay et al. 2011, Siyanova-Chanturia, Conklin and van Heuven 2011) and remembered better (Dąbrowska et al. 2009, Tremblay et al. 2011).

So there is some evidence that adults also rely on holistic processing. The question is how important a role it plays in adult processing. Most linguists seem to believe that it's quite marginal. Adults' utterances are believed to be mostly novel, and they are also thought to be longer and more complex than those produced by children. Since the number of possible combinations increases exponentially with length, it would seem that cut-and-paste couldn't possibly work for adults' linguistic abilities.

There is one practical problem with doing this kind of analysis with adults. Adults have experienced vast amounts of language; therefore, to have a reasonable sample of their experience, we need a very large corpus, ideally one containing the speech of a single individual, and such corpora are simply not available.

But we do have some very large developmental corpora — for instance, the Thomas corpus (Lieven et al. 2009), which contains 379 hours of conversation between a young child, Thomas, and his mother. In fact, it contains about 1,800,000 words produced by the mother.

The question now arises whether we can use child-directed speech as representative of adult conversational abilities? We know that adults adjust their speech when they talk to young children. However, by the time the child is four or five years old, these adjustments are negligible. There are several reasons for thinking so. First of all, most child language researchers believe that grammatical development is complete by age four or five. For instance, according to Pinker (1995), "It is safe to say that except for constructions that are rare, predominantly used in written language, or mentally taxing even to an adult (like *The horse that the elephant tickled kissed the pig*), all parts of all languages are acquired before the child turns four." Hoff (2009) makes a similar claim: "By age 5, children essentially master the sound system and grammar of their language and acquire a vocabulary of thousands of words.... The development of complex (i. e., multi-clause) sentences usually begins some time before the child's second birthday and is largely complete by age 4." And Hirsch Pasek and Golinkoff (1996: 2) assert that "It is well known, however, that children acquire most of their grammar by the time they are three years old."

So there is a general consensus that grammatical development is complete by about the age of four.

Actually this is not true. There is considerable evidence that children's grammars continue to develop throughout childhood, and well into adolescence. Most of this is as a result of exposure to written language and it can be seen as acquiring an additional register or additional set of registers over and above informal spoken discourse. But there are also effects that you can see in informal conversation. Table 1 presents some relevant data from Nippold et al. (2005), who studied the syntactic complexity of utterances produced by speakers aged from 8 to 44. They used several measures of grammatical complexity including mean length of T-unit (i. e., main clause plus any subordinate clauses attached to it) and clausal density (the number of clauses per T-unit). As can be seen from the data in the table, these continue to grow slowly but steadily from age

8 right up to 25 or so, and then growth levels off. Notice also that there is a vast amount of variation between speakers in each age group.

Table 1: Mean length of T-unit (in words) and clausal density in conversation (Nippold et al. 2005)

Age	Length of T-unit		Clausal density	
	Mean	Range	Mean	Range
8	6.74	4.42–8.44	1.18	1-1.37
11	7.31	3.67–10.56	1.25	1-1.75
13	6.88	5.56–8.52	1.17	1-1.42
17	8.33	5.83–10.32	1.30	1.08-1.56
25	9.86	6.00–13.44	1.39	1.08-1.82
44	9.56	6.88–15.16	1.38	1.12-1.77
*MOT	8.60		1.30	

The last row in the table presents the corresponding data for a sample of the mothers' speech that I am going to analyze shortly. This recording was made when Thomas was five years old, and the mother was in her mid twenties. As we can see from the table, her mean length of T-unit and clausal density are slightly below the average figure for her age, but well within the range for this group. So I conclude that her speech to Thomas is representative of adult conversational abilities.

For the purposes of this study I randomly extracted a hundred multiword utterances produced by the mother from the last transcript, which was recorded when Thomas just under five. The criteria for identifying units and deriving target utterances were the same as in the child study with one exception, namely, I also allowed two-slot frames such as *now you've VPen S* (e.g., *now you've eaten you can go out*), or *NP1 could do without NP2* (e.g., *I could do without your constant interruptions*). (In the child analysis, we only allowed frames with a single slot.)

The analysis revealed that about a third of the mother's utterances (34% to be exact) are either zero-operation utterances, in other words, exact repeats of things that are attested at least twice in the main corpus, or single operation utterances. Another 32% required 2-4 operations; 16% required between 5 and 12 operations; and 18% could not be derived. Of the 18%, 8 are lexical fails (in other words, they contain a word that is not attested at least twice in the main corpus) and 10 are syntactic fails, i.e., they cannot be derived from multiword chunks attested in the corpus.

Turning to the types of units that were used in the derivation: about 44% were invariable units (i.e., fixed phrases or single words), 48% were frames with one slot and 8% were frames with two slots. Most of the multiword units derived from the corpus are fairly short (from 1 to 4 words), which is in line with earlier research, although there were also a few units which are five, six, and seven words long.

Another interesting question that this analysis will help us address is how much processing effort we can save by using multiword chunks. The average utterance length in the sample that was analyzed in this study was 7.4 words. So if you combined each word separately, you would need 6.4 operations. The average number of operations for utterances if you use multiword chunks is 2.3. In other words, using chunks saves you about two-thirds of the processing effort.

Let us look at a few examples to see how this works in practice. One of the utterances in the test corpus was *The cover for the box is still there, it's still in the box, look.* This particular example raises several issues. First of all, is this one utterance or two — or perhaps three? For the purposes of this analysis, I followed the transcribers' intuitions about what constitutes an utterance. The Thomas corpus is transcribed in the CHAT system, and in CHAT, every utterance is transcribed on a single line. So I assumed that everything that the transcriber placed in a single line was an utterance.

To assemble this utterance, you need five chunks: *NP is still there* (which is attested 3 times in the main corpus), *the cover for the N* (attested twice), *it's still in the box* (attested twice), and *the box* and *look* (both of which are attested more than 50 times). To derive the utterance, we need to superimpose *the cover for the N* and *the box*; this gives us *the cover for the box*, which in turn is superimposed over the NP slot in *NP is still there*, giving us *the cover for the box is still there.* Then we juxtapose this expression and *it's still in the box*, and finally juxtapose *look.* This gives us *the cover for the box is still there, it's still in the box, look.*

I mentioned earlier that the frequency threshold of two tokens in the corpus may not be realistic. It is unlikely that we store everything that we hear just twice. On the other hand, bear in mind that the main corpus constitutes only about 0.5% of the mother's linguistic experience, so our best estimate of the actual frequency of the units in the mother's experience is corpus frequency multiplied by 200 — or in this case, 400.

Obviously we don't know how many times Thomas' mother actually heard or produced these expressions: it is possible that the two instances attested in the main corpus are the only ones in her experience. But it is important to note that nothing crucial hangs on the frequency assumption: if we raise the frequency threshold we can still derive the utterance — we will just need slightly more abstract chunks: *it's still in the N*, which was attested four times, instead of *it's still in the box*; *cover for NP*, attested 7 times, instead of *the cover for the N*, and *NP is still LOC*, attested 8 times, instead of *NP is still there.*

The second example that we will look at is *I don't think we really need a train track, do you?* All you need to do to assemble this rather complex utterance is to superimpose three chunks: *I don't think S do you?* (117 tokens in the main corpus), *I don't think we really need NP* (2 tokens), and *a train track* (29 tokens). Notice that this utterance could not be derived using traditional tag rules because the pronoun in the main clause does not correspond to the pronoun in the tag. But it is very frequently attested combination in the main corpus.

Our last example is *you are not going to tell me what happened when you went to Luke's house, Thomas.* This will allow us to see that even quite complex sentences involving several subordinate clauses can be assembled using preconstructed chunks. To derive the sentence, we need five units: *you are not going to VP*, which occurs 91 times in the main corpus; *tell me what happened when S*, which occurs 3 times; *you went to POSS house*, which occurs 8 times; *Luke's N*, which occurs 7 times; and *Thomas*, which occurs very often, since this is the child's name.

Again, the utterance could have been derived in different ways using different chunks. To appreciate this, let us focus on the subordinate clause *when you went to Luke's house.* Applying the rules developed by Dąbrowska and Lieven (2005), we would need to superimpose two units, *when you went to POSS house* and *Luke's N*; but there are many other possibilities, for instance:

when S + you went to POSS house (7) + Luke's N (7)

when S + you VP + NP went to POSS house (45) + Luke's N (7)

when S + you went to NP + POSS house + Luke's N (7)

when you went DIR (41) + to NP's house + Luke

when NP went to NP's house (16) + you VP + Luke

(The numbers in parentheses indicate

frequency in the main corpus; lack of a number indicates a frequency greater than 50.)

The results summarised earlier suggest that ordinary language use involves recycling chunks or what has been called "cut-and-paste". Let me emphasize again that doing cut-and-paste does require grammatical knowledge: to do it successfully, a speaker must have some understanding of constituency and know which chunk of form corresponds to which chunk of meaning; and s/he also has to know which categories or subcategories of items can be inserted in a particular slot. But, as I suggested earlier, this is different from the kind of knowledge that linguists have traditionally assumed.

I suggest that recycling chunks is the **basic mode** of sentence production: this is what speakers normally do in conversation. Is this all there is to knowing a language? Well, no. Adults do have more abstract knowledge. Some structures such as relative clauses, for example, appear to require more abstract representations. But these kinds of structures are typical for written language rather than spoken language.

To summarise: I have argued that speakers construct utterances by retrieving prefabs which partially match their communicative intentions and combining them using one of two operations, superimposition or juxtaposition. They probably try different ways of combining prefabs in parallel, and whichever utterance they come up with first wins the race, so to speak, and is actually produced. I suggest that regular combinations are available as prefabs if they are frequent enough. Unfortunately, we don't know how frequent "frequent enough" is. It is likely that there is a frequency/specificity tradeoff: more frequent units are easier to retrieve, but larger units result in simpler derivations. It is possible that different speakers resolve this conflict in different ways, so some speakers might have a preference for larger units while others may prefer smaller, more frequent units.

The possibility that different speakers have different lexically specific chunks raises some interesting questions. How can we understand each other? In what sense can two different speakers be said to speak the same language? And if language is just a collection of chunks, why are there higher-order regularities in language? I would like to conclude by briefly discussing these three issues.

First of all, how can we understand one another? As we have seen, the same utterance can be produced or interpreted using different chunks.

This means that speakers don't actually have to share exactly the same grammar to be able to communicate. It is also important to note that we are very good at guessing our interlocutors' communicative intentions. Mind-reading skills are absolutely crucial for language learning (Tomasello 2003, Tomasello et al. 2005). We would not be able to learn to speak if we weren't able to do this. When you're learning a language, or at least when you are learning your first language, you must be able to infer the meaning of the utterance from the context in which it occurred and perhaps the meanings of a few words in it; and of course adults are also very good at this (provided the context is informative enough).

Another question that arises is whether we speak the same language, or in what sense is the language that we use to communicate "the same" language. To answer it, it is useful to appeal to an analogy made by Saussure, who suggests the following: "A language, as a collective phenomenon, takes the form of a totality of imprints in everyone's brain, rather like a dictionary of which each individual has an identical copy… Thus it is something which is in each individual, but is none the less common to all." (Saussure 1972/1986: 19)

I think this is a good analogy, but it is not applied in quite right way. We all have a dictionary in our heads — but it's not the same dictionary. Although all vocabulary is shared by some speakers, it needn't be shared by all speakers. Presumably all English speakers know words like *big*, *book* or *eat*. There are also words which most educated speakers of English will know, like *malice*, *concur* and *integral*. And there are words like *heteroskedasticity and polysynthetic* which only a very few speakers share.

The same is true of larger units such as fixed phrases and frames with slots. Highly frequent units such as *I don't think so*, *I don't believe a word of it*, or *I haven't seen you for ages* are likely to be shared by everyone. (The frequencies of these in WebCorp are 369, 123, and 101, respectively.) Less frequent multiword chunks such as *What's a nice girl like you doing in a place like this?* are probably not shared by all speakers, but can still be shared by many. (There are 50 tokens of this sentence in WebCorp.)

There is also another sense in which linguistic expressions are shared, and that's the fact that communication is sharing of symbolic expressions. If I produce an utterance and you successfully interpret it, we have shared that particular linguistic expression. Notice that this inverts Chomsky's distinction between I-language and E-language. Chomsky says that we all share the same I-language, while E-language is epiphenomenal and therefore not worthy of scientific study. What I would like to suggest is that E-language is shared between speakers and we have different I-languages. And in fact there is a considerable amount of work showing that there are large individual differences in what native speakers know about the grammar of their native language, including some very basic grammatical construction (see Dąbrowska 2012 for a review): in other words, different speakers have different I-languages. So we should think of languages as something that belongs to communities rather than to individual speakers. Or you can think of individual speakers "owning" only some parts of language. As Millikan (2008) argues, it is not necessary for a linguistic convention be shared by all speakers in order for it to survive in a speech community: all that is necessary is that it succeeds in coordinating speaker-hearer communicative goals some critical proportion of the time. It doesn't have to be all of the time. It probably doesn't even have to be most of the time.

Finally, why are languages as regular as they are? Well, we all know that they are not actually that regular. There are many irregularities in language: as Sapir famously said, "All grammars leak" (1921: 39). But clearly there are patterns: a language is more than just a collection of chunks. So how can we explain these patterns? For the generativists, the answer is easy: regularities in language (as well as cross-linguistic similarities) are attributable to the fact that all humans share the same Universal Grammar. However, there is very little evidence that UG actually exists (cf. Dąbrowska under review). Another possible explanation is that speakers have a preference for general rules. Again, this does not seem to be the case. Speakers actually prefer more specific units over general schemas (see Dąbrowska 2010).

So where does regularity in language come from? There is a considerable amount of work in the functional tradition in linguistics demonstrating that many crosslinguistic patterns arise as a result of discourse pressures (Du Bois 1987, Foley and Van Valin 1984), a general preference for iconic structures (Givón 1989, Haiman 1985), and various pressures on the processing system (Hawkins 1994, Kirby 1998, Kluender 1998, Kluender and Kutas 1993). Clearly, the same factors can be used to explain the emergence of patterns within a particular language. Furthermore, speakers may approximate each other's behaviour without actually sharing the same grammar (Dąbrowska 2008a, Dąbrowska 2008b, Hurford 2000).

Finally, I would like to briefly discuss what might be a problem for models such as the one proposed here, namely, that assuming lexical storage of a large number of regular multiword chunks leads to exponential amounts of redundancy. Consider, for example, the expression *at the end of the book*. A speaker could extract a number of prefabs from it: *at the end of NP*, *at the end of the N*, *at the N of NP*, *the end*, *the end of NP*, and so on. Is this a problem? Not necessarily. Note that in order to account for the data, we don't have to assume that everybody stores everything. As I hinted earlier, it is possible that different speakers have different lexically specific units. More importantly, a grammar is not an unstructured list of expressions: it is a network of interconnected units. Some of these units are partially overlapping, so the representation of *at the end of the N* can be thought of as containing those of *at the end of NP*, *the end of NP*, *the end*, and so on: that is to say, speakers can access subchunks of a larger chunk as needed.

Another kind of objection that people sometimes make to this proposal is that the frequency effects that have been observed in the experimental studies that I mentioned earlier do not reflect retrieval but speeded-up processing: in other words, we produce and understand frequent word combinations faster because we have performed that series of computations many times before. Notice that this way of phrasing the question (retrieval or speeded-up processing?) presupposes that retrieving chunks and combining units are two distinct processes: in other words, it presupposes a dual mechanism account. But if you consider the question from the point of view of human processing, you can think of retrieval as reconstruction. This is easy to conceptualize in a connectionist model. A connectionist model learns to associate a particular input with a particular output. It is trained on a number of examples and then it is tested on some novel combinations. Is the model doing retrieval or is it doing combination? If it was exposed to a particular pattern before, we can say it's simply retrieving a previously learned combination; and if it hasn't been exposed to it, we say it's computing a novel combination. But in both cases, the model is doing exactly the same type of computation! From the model's point of view, there is no difference between the two situations.

Now consider the issue from a slightly different perspective. Suppose a second language learner is trying to remember the French expression for "what time is it?" and she can't remember it. Then the teacher says *quelle*, or *heure*, and the student says *Oh yeah*, *quelle heure est-il*? Any part of the pattern, including the intonational contour, can serve as a retrieval cue which enables the learner to complete the pattern; in this case, we are dealing with retrieval. Now a competent speaker of French knows the rules for forming questions, and they can process a stereotypical question such as *quelle heure est-il*? faster than less formulaic questions, possibly as a result of knowing the rule. But is what the competent speaker is doing *fundamentally* different from what the learner is doing? If you assume that lexical items and constructional schemas of varying degrees of specificity are part of the same network of symbolic units in a content-addressable memory, there is no fundamental difference between the two. There is a difference in that the learner might not know the general schema for forming questions, but this is not a fundamental difference.

The final point that I would like to make is that if we want to do cognitive linguistics, we should remember that language is a *cognitive* phenomenon. This means that we must seriously think about cognitive constraints on how languages are learned and processed, and that we should look to work on human cognition for solutions to linguistic problems — to work on human pattern-finding skills, mind-reading skills, and the fact that we have a content addressable long-term memory with a vast capacity.

References:

[1] Dąbrowska, E. (2008). The effects of frequency and neighbourhood density on adult speakers' productivity with Polish case inflections: An empirical test of usage-based approaches to morphology. *Journal of Memory and Language*, 58, 931 – 951.

[2] Dąbrowska, E. (2008). The later development of an early-emerging system: The curious case of the Polish genitive. *Linguistics*, 46, 629 – 650.

[3] Dąbrowska, E. (2010). The mean lean grammar machine meets the human mind: Empirical investigations of the mental status of rules. In H.-J. Schmid & S. Handl (Eds.), *Cognitive foundations of linguistic usage patterns. Empirical approaches* (pp. 151-170). Berlin: Mouton de Gruyter.

[4] Dąbrowska, E. (2012). Different speakers, different grammars: Individual differences in native language attainment. *Linguistic Approaches to Bilingualism*, 2, 219 – 253.

[5] Dąbrowska, E. (under review). What exactly is Universal Grammar, and has anyone seen it?

[6] Dąbrowska, E., & Lieven, E. (2005). Towards a lexically specific grammar of children's question constructions. *Cognitive Linguistics*, 16, 437 – 474.

[7] Dąbrowska, E., Rowland, C., & Theakston, A. (2009). The acquisition of questions with long-distance dependencies. *Cognitive Linguistics*, 20, 571 – 596.

[8] Du Bois, J. (1987). The discourse basis of ergativity. *Language*, 63, 805 – 855.

[9] Foley, W. A., & Van Valin, R. D. (1984). *Functional Syntax and Universal Grammar.* Cambridge: Cambridge University Press.

[10] Givón, T. (1989). *Mind, Code and Context: Essays in Pragmatics.* Hillsdale, NJ: Lawrence Erlbaum.

[11] Gurevich, O., Johnson, M. A., & Goldberg, A. E. (2010). Incidental verbatim memory for language. *Language and Cognition*, 2, 45 – 78.

[12] Haiman, J. (1985). *Iconicity in Syntax.* Cambridge: Cambridge University Press.

[13] Hawkins, J. A. (1994). *A Performance Theory of Order and Constituency.* Cambridge: Cambridge University Press.

[14] Hirsh-Pasek, K., & Golinkoff, R. M. (1996). *The Origins of Grammar: Evidence from Early Language Comprehension.* Cambridge, MA: MIT Press.

[15] Hoff, E. (2009). Language development at an early age: Learning mechanisms and outcomes from birth to five years. In *Encyclopedia of Early Childhood Development.*

[16] Hurford, J. R. (2000). Social transmission favours linguistic generalisation. In C. Knight, M. Studdert-Kennedy & J. R. Hurford (Eds.), *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form* (pp. 324-352). Cambridge: Cambridge University Press.

[17] Kirby, S. (1998). Fitness and the selective adaptation of language. In J. R. Hurford, M. Studdert-Kennedy & C. Knight (Eds.), *Approaches to the Evolution of Language: Social and Cognitive Bases* (pp. 359-383). Cambridge:

Cambridge University Press.
[18] Kluender, R. (1998). On the distinction between strong and weak islands: A processing perspective. In P. W. Culicover & L. McNally (Eds.), *Syntax and Semantics* 29: *The Limits of Syntax* (pp. 241-279). San Diego: Academic Press.
[19] Kluender, R., & Kutas, M. (1993). Subjacency as a processing phenomenon. *Language and Cognitive Processes*, 8(4), 573 – 633.
[20] Lieven, E., Salomo, D., & Tomasello, M. (2009). Two-year-old children's production of multiword utterances: A usage-based analysis. *Cognitive Linguistics*, 20, 481 – 508.
[21] Millikan, R. G. (2008). A difference of some consequence between conventions and rules. *Topoi*, 27, 87 – 99.
[22] Nippold, M. A., Hesketh, L. J., Duthie, J. K., & Mansfield, T. C. (2005). Conversational v. expository discourse: A study of syntactic development in children, adolescents and adults. *Journal of Speech, Language and Hearing Research*, 48, 1048 – 1064.
[23] Pinker, S. (1995). Language Acquisition. In L. R. Gleitman & M. Liberman (Eds.), *An Invitation to Cognitive Science* (2 ed., Vol. 1: Language). Cambridge, MA: MIT Press.
[24] Pollack, I., & Pickett, J. M. (1964). Intelligibility of excerpts from fluent speech: auditory vs. structural context. *Journal of Verbal Learning and Verbal Behavior*, 3, 79 – 84.
[25] Sapir, E. (1921). *Language: An Introduction to the Study of Speech.* New York: Harcourt, Brace and Company.
[26] Saussure, F. d. (1972/1986). *Course in General Lingusitics* (R. Harris, Trans.). Peru, Illinois: Open Court Publishing Company.
[27] 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 multi-word sequences. *Experimental Psychology: Learning, Memory, and Cognition*, 37, 776 – 784.
[28] Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Child Language Acquisition.* Cambridge, MA: Harvard University Press.
[29] Tomasello, M. (2006). Construction grammar for kids. *Constructions*, *SV*1-11, 1 – 23.
[30] Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28, 675 – 691.
[31] Tremblay, A., Derwing, B., Libben, G., & Westbury, C. (2011). Processing advantages of lexical bundles: Evidence from self-paced reading and sentence recall tasks. *Language Learning*, 61, 569 – 613.

Received: 2012 – 11 – 03

Bionotes: Ewa Dąbrowska's research interests include cognitive linguistics, language acquisition, the mental status of rules, and individual differences in linguistic knowledge. She has published monographs on the semantics of case (*Cognitive Semantics and the Polish Dative*, Mouton de Gruyter, 1997) and the relationship between linguistics and other disciplines studying language (*Language, Mind and Brain: Some Psychological and Neurological Constraints on Theories of Grammar*, Edinburgh University Press/Georgetown University Press, 2004), an edited volume on cognitive linguistic approaches to language acquisition (with Wojciech Kubiński, Universitas, 2003), and numerous articles. She is the editor of *Cognitive Linguistics*, Vice President of the UK Cognitive Linguistics Association and a member of the Governing Board of the International Cognitive Linguistics Association.