The usage-based model

Agata Kochańska

Warsaw University

Fundamental questions:

➤What shape does our knowledge of a language take?

➤How is this knowledge acquired?

➤How is this knowledge put to use?

Generative Grammar

➢ The shape that our knowledge of a language takes

autonomy of language thesis: language viewed as a separate module of the human mind, which is governed by its own principles

modularity of language and **autonomy of syntax**: grammar (rules) and lexicon (a list of idiosyncratic information); syntax independent of semantics and phonology

the grammar of a language **generates** all and only grammatical sentences of this language

consequence: regular complex expressions are semantically **fully compositional**

emphasis on the economy of description – whenever possible, human linguistic knowledge is represented in the form of **maximally general abstract algorithms**; less general representations are postulated only when necessary

grammar: a set of highly abstract rules operating on abstract (purely syntactic) symbols; deep structures which are never empirically observable and which are often substantially different from surface structures

➢First language acquisition in GG

poverty-of-the-stimulus argument: children cannot acquire a language via mere induction from linguistic input they receive; to account for the fact that the do acquire a language we need to postulate inborn grammatical knowledge, the knowledge of the rules of Universal Grammar (**Innateness Hypothesis**)

UG is the set syntactic rules universal for all languages; the knowledge of the rules of UG is inborn; these rules contain **parameters**, whose values for a particular language has to be set in the process of acquisition; the rules of UG with parameters set for a particular language constitute the **core grammar** of that language

acquisition of **core grammar**: setting appropriate values to the parameters within the rules of UG

the **peripheral grammar**: the set of grammatical rules specific for a particular language

acquisition of the **peripheral grammar**: the rules of peripheral grammar have to be actually acquired) from the linguistic input

regular aspects of the periphery are acquired via the mechanism of extracting maximally general rules (cf. e.g. the mechanism of acquiring the pattern of forming plural nouns or the past tense forms of verbs in English)

irregular aspects of the periphery are acquired via memorization (cf. e.g. *sink – sank, drink – drank*)

two mechanisms of language acquisition (**dual-mechanism** theory)

- rote-learning – induction from the input and memorization responsible for the acquisition of the lexicon

- the abstraction of maximally general rules via deductive learning – responsible for the acquisition of regular grammatical patterns

Selected consequences of the generative model of first language acquisition:

- General inductive learning mechanisms play a relatively minor role in the acquisition process

- In the acquisition process children acquire **either** general grammatical rules **or** idiosyncratic lexical information; these two aspects of linguistic knowledge are acquired via different learning mechanisms (deductive **rule abstraction** vs. inductive **rote-learning**) and are represented in a different form in the mind (general **rule** vs. specific **instance**)

•**Continuity assumption**:

since UG is an inborn knowledge present in the mind of both a young child acquiring a language and a mature speaker of a language, it has to be assumed that a young child's linguistic representations are **of the same fundamental nature** as an adult's representations (since it is assumed that the grammatical knowledge of an adult speaker resides in maximally general abstract algorithms, it also has to be assumed that the inborn grammatical knowledge of a young learner takes the form of maximally general abstract algorithms

Agata Kochańska

Warsaw University

Logical argument against the GG acquisition model:

Based on Dąbrowska (2000a)

Chomskyan child – a poor inductive learner with vast inborn knowledge

the problem of the acquisition of inflectional morphology

> inflectional morphology is language-specific and hence, has to be acquired

GG, at least in its origins, seems to be an English-centered theory of language

English is quite exotic, as far as its inflectional morphology is concerned

its inflectional morphology is practically non-existent

Whatever inflectional morphology it has, this morphology is prevalently regular (conforming to a single general pattern), with just a handful of exceptions

Some of the things that a child acquiring the Polish case system has to find out:

➢ What case categories of Polish are (how many case distinctions the language makes)
   common pattern – 2 to 7

   Finnish – 15   Avar (a Caucasian lg) – 27 Tabassaran (a Caucasian lg) – 53

   ➢ What the formal markers of case categories are
      Case-like notions may be conveyed by:
      affixes attached to a noun, stem alternations, prepositions, verbal affixes,
      various combinations of these devices, word order

➢Where to put the case marker

  A case affix may be attached to a noun stem, it may follow other affixes, there may be different orderings with different cases.
  Case marking may be attached to the determiner (cf. German) or the entire NP (cf. Saxon Genitive in English)

➢one case may have multiple exponents

  a. *studenta*$_{GEN.}$ (student)   -a
  b. *makaronu*$_{GEN.}$ (pasta)   -u
  c. *kobiety*$_{GEN.}$ (woman)   -y

➢The same marking may be an exponent of different cases (syncretism)

  a. *pokój studenta*$_{GEN}$ ('student's room')   -a   GEN.
  b. *Spotkałam studenta*$_{ACC.}$ ('I met a student)   -a   ACC.

➢Since the same case has different exponents with different nouns, the child has to establish what morphological classes Polish nouns are divided into

The total number of distinct paradigms for Polish nouns is 52

Class membership depends on a number of factors including gender, morphological properties of the stem and the nominative, presence of certain derivational endings (e.g. diminutives), and – to a certain extent – on semantic properties

➢The addition of a case ending may induce stem alternations including both vowel and consonant mutations

a. *pies*$_{NOM.}$ – *psu*$_{DAT.}$ ('dog')   b. *kobieta*$_{NOM.}$ – *kobiecie*$_{DAT.}$ ('woman')

➢What the semantic functions of each case are

a. *Piotr dał Ewie*$_{DAT.}$ *książkę* (Piotr gave Ewa a book)

b. *Patrzył Ewie*$_{DAT.}$ *na nogi* (Piotr stared Ewa at legs 'Piotr stared at Ewa's legs')

c. ?*Patrzył Ewie*$_{DAT.}$ *na lampę* (Piotr stared Ewa at lamp 'Piotr stared at Ewa's lamp')

➢The distribution of cases with numerals

a. *dwa komputery*$_{NOM.}$ (two computers)

b. *pięć komputerów*$_{GEN.}$ (five computers)

c. *dwadzieścia dwa komputery*$_{NOM.}$ (twenty two computers)

d. dwadzieścia pięć *komputerów*$_{GEN.}$ (twenty five computers)

e. *paczka z dwoma komputerami*$_{INSTR.}$ (a parcel with two computers)

f. *paczka z pięcioma *komputerów*$_{GEN.}$/*komputerami*$_{INSTR.}$ (a parcel with five computers)

How do children acquire the Polish case system?

➢ All six case categories appear before the age of two

➢ Most case endings are used correctly from the moment they emerge

➢ Children can extract general patterns from the input, as witnessed by the occurrence of overgeneralization errors (in the third year of life) and the correct inflection of nonsense words (preschool children)

➢ The acquisition of cases in numeral + noun constructions: errors consisting in the incorrect choice of case are rare (children between 1;9 and 5;0)

Conclusion to the logical argument:

Since Polish-speaking children successfully acquire a very complex system of Polish inflectional morphology, they must be equipped with powerful mechanisms of inductive learning.

This does not prove that there is no inborn linguistic knowledge.

However, this fact casts doubts on the validity of the poverty-of-the-stimulus argument.

children are equipped with powerful mechanisms of inductive learning

the success in the acquisition process may be attributable precisely to those mechanisms, rather than to any inborn linguistic knowledge

Cognitive research on language acquisition

Major topics for empirical research on first language acquisition

•to identify the nature of psycholinguistic units with which children operate (the nature of mental representations underlying linguistic performance at particular stages of development)

GG: **continuity assumption** (the mental representations constituting the knowledge of a language of a child are of the same basic nature as the mental representations constituting the knowledge of a language of an adult speaker); **not** an **empirical finding** but a result of *a priori* **theoretical considerations** (Innateness Hypothesis)

Cognitive approach: what mental representations lie behind particular observable linguistic behaviors of a child is an **empirical question** (it should not be resolved on a purely theoretical a priori grounds)

•to identify the nature of mechanisms via which particular kinds of representations emerge at particular stages of development

GG: Regular grammatical patterns are acquired via deductive abstraction of a maximally general rule

Lexical knowledge is acquired via induction from the input and memorization

Cognitive approach: the nature of the relevant mechanisms is again an **empirical issue**

Empirical research:

Tomasello (2000)

➢a study of young Italian-speaking children

Pizutto and Caselli (1992 and 1994)

six possible person-number forms of verbs in the present tense in Italian

about **half** of all the **verbs** used by each of the children under investigation were used in **one form** only, and an additional **40%** were used with **two** or **three** forms

of the **ten percent** of verbs that were used in four or more forms, approximately **half** were **highly frequent**, **highly irregular** forms that could only have been **learned by rote** (they couldn't be derived by an application of an abstract rule)

Tentative conclusion:

Young children are not fully productive in their use of person-number verbal endings

Hence, they don't seem to operate with fully general rules
(e.g. $V_{PRES.\,3\,SG.} \rightarrow V + X$)

Instead, they seem to operate with generalizations which are partially lexically specific
(e.g. $drink_{PRES.\,3\,SG.} \rightarrow drink + X$)

➢The overgeneralization experiment with English-speaking children

rooks *et al.* (1999)
the experiment aimed at inducing overgeneralizations with respect to existing words

subjects: children between 3.5 to 8 years of age

words used in the experiment – English verbs with fixed transitivity (such as e.g. *disappear* – exclusively intransitive and *hit* – exclusively transitive)

four pairs of synonymous verbs: one member of a pair was a verb learned early by children and used frequently by adults, while the other was learned later by children and was used less frequently by adults (*come – arrive*)

experimenters asked questions aimed at inducing overgeneralization (transitive use of intransitive verbs)

Finding:

Children of all ages were less likely to overgeneralize the verbs which were learned earlier by children and used more frequently by adults

Tentative conclusion:

Words such as *come* are much more frequently used than words such as *arrive*

Hence, children can master the patterns connected with *come* much more thoroughly (these patterns become much more strongly automatized / routinized / **entrenched**)

The level of **entrenchment** seems to have real psychological consequences – the entrenched patterns are much stronger, more easily activated

➢A case study of the linguistic development of questions in the spontaneous speech of a single English-speaking child, Naomi

Dąbrowska (2000b)

Question:

the nature of linguistic representations in the child's mind in the process of language acquisition (more specifically, the nature of mental representations underlying an English speaking child's ability to form questions)

standard generative view: abstract rules of subject-auxiliary inversion and *wh*-movement

Comment on the methodology of a case study

disadvantages: the problem of generalizability

advantages: a detailed case study of a single child makes possible tracing the development of this in all its specificity

GG on the issue of linguistic productivity:

"Except for a ridiculously small number (…) all actual sentences are of a probability so low as to be effectively zero and the same is true of structures. (…) In general, it is a mistake to assume that – past the earliest stages – much of what the child acquires is acquired by imitation. This could not be true on the level of sentence formation, since most of what the child hears is new and most of what he produces, past the very earliest stages, is new."

Chomsky (1964:37)

Methodology:

detailed case study of the linguistic development of questions in the spontaneous speech of a single child, Naomi

data: transcripts of about 90 hours of Naomi's spontaneous conversation in a home setting; the span of time from 1;6.16 to 3;8.19

all utterances produced by Naomi which ended in a question mark in transcripts were electronically extracted from the corpus

Then they were coded by hand
(all relevant aspects of the grammatical structure of these utterances were tagged by hand in the corpus of extracted transcripts, so that the relevant structures could be easily searched for)

the study was concerned with syntactic questions
(utterances which were at least two words long and which begin with either a *wh*-word or an auxiliary and end in a question mark)

1439 questions produced by Naomi (out of 2189) met these criteria

Specific question:
What proportion of Naomi's questions are formulaic?

formulaic - either "big words" memorized (rote-learned) by child in their entirety or "big words" with a single slot in which variation is possible

the point: is to establish the role of imitation / rote-learning in the acquisition of questions

Problem: it is difficult to determine whether a particular utterance

    was retrieved from memory

    was produced by manipulating the material retrieved from memory

    was assembled from scratch

Working strategy:

combinations of words that recur again and again are likely to be stored

any sequence of simple units which occurs at least **5 times** in Naomi's corpus is regarded as **formulaic**

any sequence of simple units which recurs at least **10 times** in Naomi's corpus is regarded to be a **major formula**

Examples of Naomi's formulas:

Do – (you) – want – THING?

Is – THING *– going to work?*

Can – (I) – PROCESS?

Would – you – PROCESS?

What – ('s – THING) *doing?*

What's – Naomi – do?

What – (s – THING) *– got?*

Where – ('s) – THING *– go?*

Formulaicity in Naomi's questions

age	major formulas	minor formulas	non-formulaic	total questions	% formulaic
1;6-1;11	553	13	31	597	94.8
2;0-2;5	430	25	44	499	91.2
2;6-2;11	146	31	51	228	77.6
3;0-3;8	61	12	42	115	63.5
Total	1190	81	168	1439	88.3

Preliminary discussion:

The **vast majority** of Naomi's questions are **formulaic** by the adopted criterion

However:

The fact that the vast majority of Naomi's questions are formulaic by the statistical criterion does not mean that they are in fact retrieved from memory (rather than generated by rules on each specific occasion)

Another possibility:

Naomi did in fact have a fully developed grammar capable of generating any interrogative structure
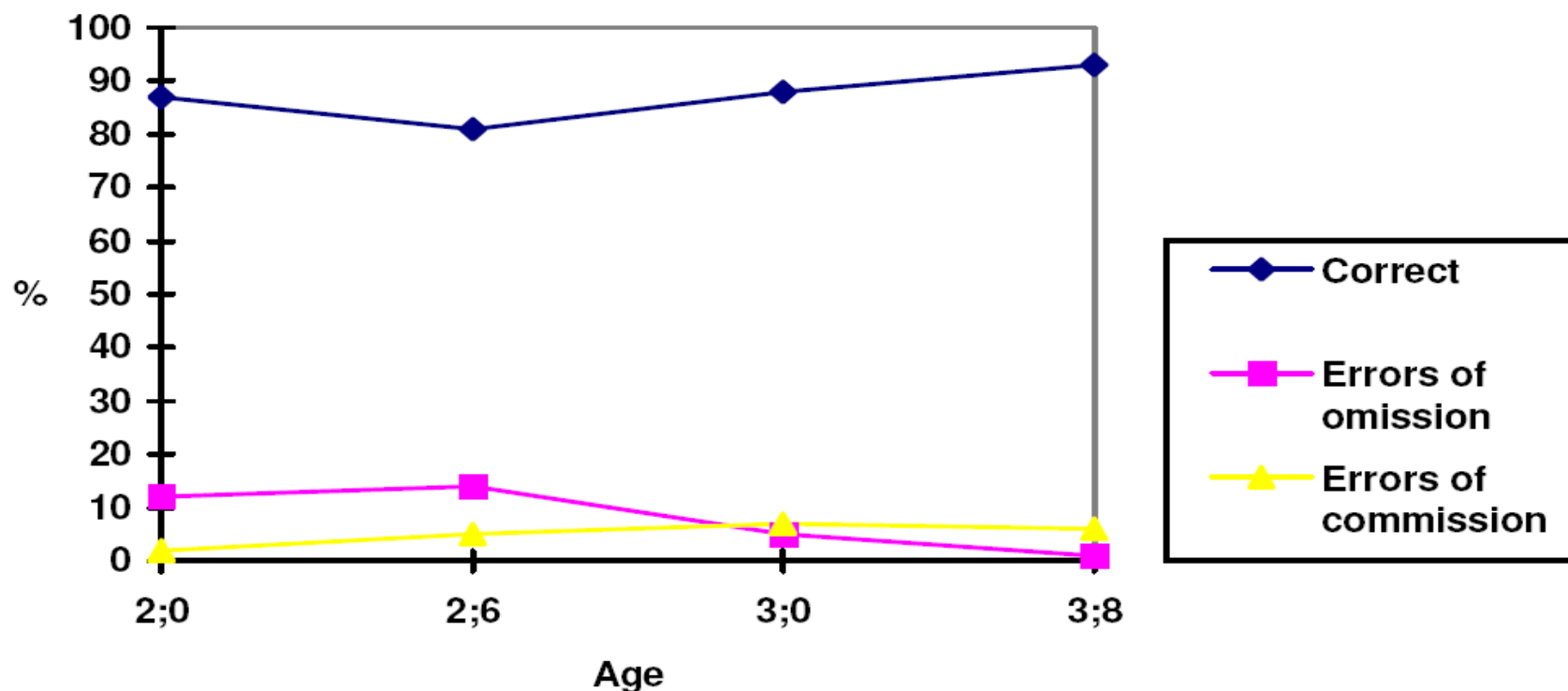
But:

She chose to make a very limited use of this grammar

Further evidence

➢level of accuracy in Naomi's performance

no sudden rise in the level of grammatical accuracy (this could signal the acquisition of the rule)

Naomi is pretty accurate from the very start

the development of specific formulas – the permission-seeking question

Can I get down?	[repeated four times]	1;11.9
Can I get up?		1;11.9
Can eat it ice cream?		1;11.16
Can lie down?	[repeated twice]	1;11.16
Can do this?		1;11.16

All these early questions with *can* are extremely **stereotypical**:

➤ variability only in the slot PROCESS

➤ the auxiliary is always *can*

➤ it is always placed at the beginning of the sentence (no "uninverted" questions)

➤ although the first person pronoun is often left out; the understood agent of the action is always Naomi herself

Could do this?   1;11.21
Could eat that?   1;11.21
Could I throw that?   2;0.3

As the formula gets analyzed, it becomes more and more flexible

The next slot where variability becomes possible is the auxiliary slot (*can – could* variability)

Can you draw eyes?   2;0.28
Please can we do that?   2;8.14

Finally, variability appears in the subject slot

Can – I – get – down?
Can – I – PROCESS*?*
ABILITY VERB – *I –* PROCESS*?*
ABILITY VERB – PERSON – PROCESS?

progression from **non-flexibility** to **flexibility**, rather than from **non-accuracy** to **accuracy**

such a progression could be characterized as a progression from **formula** to **schema**

If Naomi's ability to form questions really develops in a piecemeal fashion from rote-learned (memorized) **holophrases/formulas** via a series of **progressive schematizations** resulting from **abstracting commonalities** observable in the data, such a development requires the occurrence of three processes:

➢ phonetic segmentation of holophrases/formulas

➢ semantic analysis of the formula

➢ establishing links between the component chunks of phonetic material and the component chunks of meaning

Naomi's early usage is highly stereotypical/rigid and only gradually develops towards greater flexibility.

This does not by itself prove that Naomi's ability to form questions develops as the result of analysis and schematization of formulas

"It is also possible that the more mature usage seen in the later transcripts is attributable to a different mechanism (abstract combinatorial rules) and that the use of formulas is no more than a communication strategy which complements rules at a time when they are not yet fully developed"

(Dąbrowska 2000b:96)

Evidence for formula-centered learning

➢ The first recorded uses of a formula are often more complex than subsequent uses

Where'd it go?   before   *Where go?*
Is this doggie?   *Is toaster?*
  What's this? at 11;10.3   *What?* (twelve weeks later)
  Where's Daddy?   *Where is Daddy?*

➤Stereotyped usage

- questions with *can /could* had first-person subjects and the auxiliary was consistently placed first

- all other questions with the first-person subject are uninverted

- with a single exception all questions with *will* and *would* are inverted and have the second-person subject

- of the 31 questions with the auxiliary *do* 29 have the pronoun *you* as subject and 22 of 29 have either *want* or *wanna* as the main verb

The usage-based model

non-autonomy of language: human basic cognitive abilities are also at work in linguistic functioning

➤ **entrenchment**

whatever we do, from the simplest motor action to a highly complex mental activity, consists – at least in part – in executing some patterns of neural firing (**cognitive event**)

apparently, each such execution leaves in the neural cells some kind of **neurochemical "trace"**, which facilitates subsequent execution of the same pattern – this is the basis of the mechanism of learning that we call **routinization** / **automatization** / entrenchment

a cognitive event receives a **unit status** when its execution becomes automatized and starts constituting an established **routine**

when a complex cognitive event becomes routinized, it can be carried out without paying conscious attention to how its component parts should be integrated into the complex whole

the whole event comes as a "**prepackaged assembly**" that can be manipulated as an integrated whole (cf. e.g. Langacker 1987:100)

the mechanism of routinization / automatization / entrenchment is a general learning mechanism operative e.g. in learning how to maintain balance, walk, grasp objects, ride a bike, etc.

entrenchment is a result of **frequent recurrence**

➢**abstraction**

there is evidence that the brain uses overlapping populations of neurons to represent similar conceptualizations

the schema is the part of the representation which is shared by several instances; it is already implicit in the first instance acquired by the learner

as new instances are added to the learner's repertoire, the schema becomes more and more entrenched (commonalities get entrenched; differences are filtered out)

schema extraction is not a result of the process of seeking for generalizations, but rather "a by-product of the way information is organized in long-term memory" (Dąbrowska 2000b:99)

➢**categorization** (by **prototype** and by **schema**)

➢**composition**

The emergence of linguistic knowledge

Linguistic structure emerges from actual usage events through the processes of categorization, entrenchment and abstraction, as well as composition

> **usage event:** "the pairing of a vocalization, in all its specificity, with a conceptualization representing its full contextual understanding" (Langacker 2000:9)

human linguistic knowledge emerges through abstracting commonalities observable in actual uses of linguistic expressions with specific phonetic shapes and specific, context-bound meanings.

The process of acquiring a language resides in the progressive **entrenchment** of frequently encountered structures and the **abstraction** of progressively more and more schematic mental representations that grasp the regularities observed in actual language use.

With repeated use, this process brings about the gradual **entrenchment** of those aspects of form and meaning that are recurrent in multiple uses, as well as the filtering of those elements that do not recur.

The degree of entrenchment is a function of the **frequency** of occurrence

high **token frequency**: results in strong entrenchment of a specific expression

high **type frequency**: results in extraction and entrenchment of a pattern

"Expressions are learned by being encountered on multiple occasions, engendering contextual understandings that are similar in certain respects and diverge in others. Consistently recurrent features of these understandings are reinforced and progressively 'entrenched', whereas features that do not recur simply 'cancel out' and fail to achieve conventional symbolic association with the form. 'Linguistic knowledge' resides in structures that become cognitively entrenched and achieve the status of conventional units. Relative to the usage events giving rise to them, such units are necessarily **schematic** (i.e. characterized in lesser specificity and detail). However, any facets of the context that recur across the supporting usage events will tend to be retained as specifications of the abstracted conventional units"

(Langacker 1997:236; boldface used by the author)
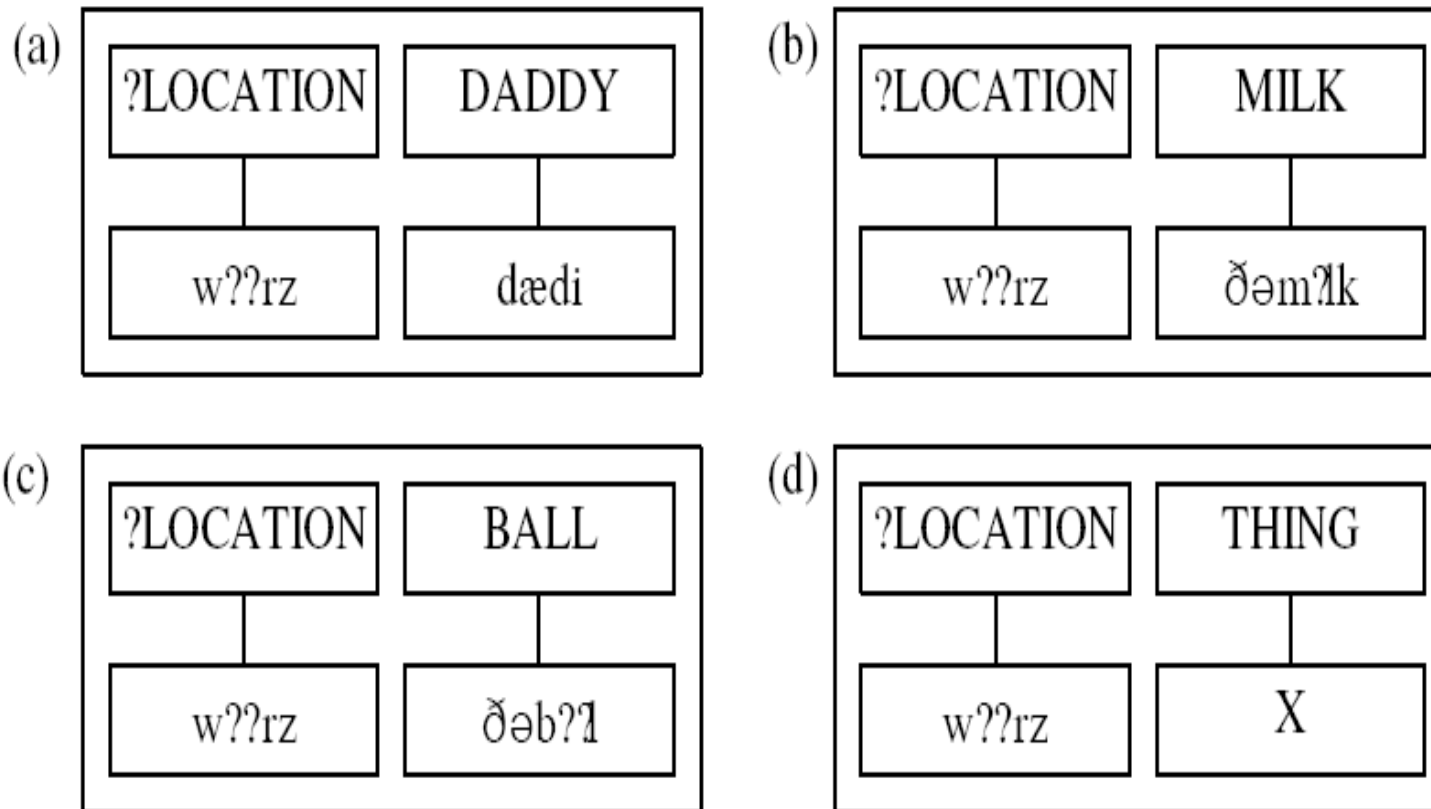
tacit knowledge of grammatical patterns: grammatical patterns (constructional schemas) are not separately represented; instead, they are inherent in their instantiations

Language system

➢ inventory of conventional linguistic resources comprising units which range from specific expressions to schemas (continuum from specific to schematic), as well as from individual morphemes to complex expressions (continuum from simple to complex); massive redundancy; massive polysemy

➢ both lexical and grammatical resources are symbolic units (**content requirement**)

➢ lexicon and grammar are not two separate components of language with very different properties

   instead, they should be viewed as forming a continuum of symbolic resources

➢ **partial compositionality** of meaning

➢the **non-constructive** nature of **grammar** (linguistic convention):

"It is not the linguistic system *per se* that constructs and understands novel expressions, but rather the language user, who marshals for this purpose the full panoply of available resources. In addition to linguistic units, these resources include such factors as memory, planning, problem-solving ability, general knowledge, short- and longer-term goals, as well as full apprehension of the physical, social, cultural and linguistic context"

(Langacker 2000:9)

the role of linguistic convention is to serve as hints in constructing and understanding expressions in actual usage-event

conventional linguistic units are invoked as the **standard** relative to which the degre of well-formedness of the relevant usage-event (**target**) is assessed (**full** vs. **partial sanction**)

Our knowledge of a language is a vast and massively redundant network, within which the same aspects of linguistic knowledge are represented as specific memorized expressions, as low-level patterns, and as patterns of progressively greater generality

Question:

Within this vast network, how is the relevant node selected as **the standard** sanctioning the relevant usage-event?

interactive activation model

"By virtue of gross similarities or shared components any facet of the input is presumed capable of activating a variety of established routines as possible standards of comparison. The candidate routines run in parallel; their specifications are continuously matched against those of the target, and constitute "expectations" about its further more specific properties. As the expectations of a given routine are successively satisfied, its level of activation is progressively augmented, and beyond a certain level it inhibits the other candidates. The system thus converges on a "solution", whereby the routine that most closely matches the target eclipses other alternatives and emerges as the primary categorizing structure" (Langacker 1987:429)

Two factors influence the selection of the categorizing structure (active node):

➤the cognitive salience of possible candidates

    inherent likelihood of activation (resulting from the degree of entrenchment)

➤their elaborative distance vis-à-vis the target

the two factors are often antagonistic

      went   vs. strived

low-level schemas are more frequently activated as sanctioning units than high-level general patterns (likelihood of activation as a function of salience and the degree of overlap with the usage-event to be sanctioned)

References:

rooks, Patricia, Michael Tomasello, Lawrence Lewis, and Kelly Dodson.

999. "Children's overgeneralization of fixed transitivity verbs: The entrenchment hypothesis". *Child Development 70,* 1325-1337.

homsky, Noam.

964. "Formal discussion of Miller and Ervin". In Bellugi, Ursula and Roger Brown (eds.). *The Acquisition of Language. Monographs of the Society for Research in Child Development 29.* (Serial no. 92). 35-42.

abrowska, Ewa

2000b "From a Chomskyan child: Is there a Polish? The English question" for language learning". In Perkins, M.R. and S. Howard (eds.). *New Directions in Language Development and Disorders*. New York: Plenum Publishers. 85-96.

Langacker, Ronald W.

1987. *Foundations of Cognitive Grammar.* Vol 1. *Theoretical Prerequisites*. Stanford: Stanford University Press.

1991. *Foundations of Cognitive Grammar.* Vol 2. *Descriptive Application*. Stanford: Stanford University Press.

1997. "The contextual basis of cognitive semantics". In Nuyts, Jan and Eric Pedersen (eds.). *Language and conceptualization*. Cambridge: Cambridge University Press. 229-252.

2000. "A dynamic usage-based model". In Barlow, Michael and Suzanne Kemmer (eds.). *Usage-Based Models of Language.Stanford: CSLI Publications*. 1-63.

Pizutto, Elena and Christina Caselli

1992. The acquisition of Italian morphology. *Journal of Child Language 19,* 491-557.

1994 The acquisition of Italian verb morphology in a cross-linguistic perspective. In Levy, Y. (ed.). *Other Children, Other Languages. Hillsdale,*NJ: Erlbaum.

Tomasello, Michael.

2000. First steps towards a usage based theory of language acquisition. *Cognitive Linguistics* 11-1/2, 61-82.

a.kochanska@uw.edu.pl

Thank you !