Bootstrapping lexical and syntactic acquisition.

Anne Christophe*
Séverine Millotte°.

*Laboratoire de Sciences Cognitives et Psycholinguistique, EHESS/CNRS/DEC-ENS, Paris
and Maternité Port-Royal, AP-HP, Université Paris Descartes
°Laboratoire de Psycholinguistique Expérimentale, Université de Genève

Lisbon
10th March 2008
Bootstrapping problems

- Bootstrapping problems
- Phonology
- Lexicon
- Syntax
- Semantics
- Function words
- Phonology
- Acoustics
Bootstrapping problems

Phonological bootstrapping
(Morgan & Demuth, 1996)
Model of processing and acquisition

Syntactic processing

[the xxx]_{NP} [is xx]_{VP} [an xx]_{NP}

Function words

[ṭā om liṭ on lb ī j]_{PP} [Isiti]_{PP}

[ən 1p əm l]_{PP}

Phonological and prosodic representation

Speech signal

"the little boy is eating an apple"
Phonological phrase boundaries constrain lexical access.

"the little boy is eating an apple"
Phonological phrase boundaries constrain lexical access: Word detection.

American 13-month-olds.

Method:
variant of conditioned head-turning.

[The church] [with the most paper spires] [is heavenly].
[The man] [with the least pay] [perspires constantly].

Gout, Christophe & Morgan (2004) *Journal of Memory and Language*
Word detection: French 16-month-olds

[La rangée de *balcons*] [fait face au cloître] [du monastère]
[La grande salle de *bal*] [*confère* un air solennel] [au château].

Phonological phrase boundaries constrain lexical access … and syntactic processing

- a phonological phrase boundary is interpreted on-line as a word boundary.

"The man with the least pay perspires constantly"
Use of phonological phrases in on-line syntactic processing (adults)

- Locally ambiguous sentences

  - Verb:

    [le petit chien]\_NP [mord la laisse]\_VP [qui le retient]…

    (the little dog bites the leash that restrains it)

  - Adjective:

    [le petit chien mort]\_NP [sera enterré demain]\_VP…

    (the little dead dog will be buried tomorrow…)

→ Sentence completion task:

  listen to sentence beginning, freely complete sentence
Conclusion: when prosodic cues are well-marked, adults exploit them to constrain their syntactic analysis of sentences.

- are these cues exploited on-line?
- are they produced spontaneously by naïve speakers?

Are prosodic cues exploited on-line? Yes

Task: abstract word detection; e.g. ’mordre’ (to bite) respond to verb sentences, refrain from responding to adjective sentences.

Results: fast responses only (given at the end of the ambiguous word)

Are prosodic cues spontaneously produced by naïve speakers? Yes

Six naïve speakers produced the ambiguous sentences: they are perceived as unambiguous by listeners…

Phonological phrase boundaries constrain lexical access and syntactic processing.

"the little boy is eating an apple"
Special role for function words

- **Syntactic processing**
  - [the xxx]_{NP} [is xx]_{VP} [an xx]_{NP}

- **Function words**
  - [the little boy is eating an apple]

- **Phonological and prosodic representation**
  - "the little boy is eating an apple"

- **Speech signal**
Function words:


Function words:

- Can be used to categorize content words:
  - ‘je jaurime’ -> ‘jaurime’ is a verb, refers probably to an action ('it blicks')
  - ‘la jaurime’ -> ‘jaurime’ is a noun, refers probably to an object ('the blick')

Höhle, Weissenborn, J. et al. (2004). *Infancy*, 16-month-old German infants: an article predicts a noun (although a pronoun does not yet predict a verb)

- Potential problems:
  - 'je la mange' (*I eat it*), la+X but X is not a noun…
  - more generally, just looking at the preceding word may not be enough to determine the syntactic category of a word (e.g. 'you stupid boy', *stupid* is not a verb)
Grammatical categories (noun, verbs) from distributional analyses

- Frequent Frames: Mintz, *Cognition*, 2003
- Frame: Ordered pair of words with one intervening word
- “we have to *put it in the bag*”

Frame: to__it tally frames, keep the most frequent ones (e.g. you__it, you__to, you__the, the__ is, ...)
Examples of frames: (Mintz, 2003)

- the _ is: 79 tokens, 47 types, 100% nouns
  
  bag  barn  basket  bear
  book  box  boy  bug
  bus  car  couch  cow
  cream  dinosaur  dog  doggy

- you _ it: 433 tokens, 93 types, 100% verbs
  
  put  see  do  did
  want  fix  turned  get
  got  turn  throw  closed
  think  leave  take  open
Does it work in other languages?

• Potential problems in French:
  • homophony between articles le/la/les and object clitics le/la/les
    'la + N' as in 'la poire' the pear
    'la + V' as in 'je la mange' I eat it
  • inverted word order for object clitic so that some efficient frames in English (You_it) don't translate; (You want it 'tu le veux')
  • more varied function words (gender/number distinctions; 'le la les' the).
Frequent frames in French and English

Frames:
- \([A \times B]\)
- Fronting context:
  - \([A \ B \ x]\)
  - (‘je le veux’ / I want it)
- Back context:
  - \([x \ A \ B]\)

Discontinuity.

Chemla, Mintz, Bernal & Christophe (2008) Developmental Science
Function words and categorization:

- Function words may support reliable categorization (even with a relatively simple-minded analysis, such as frequent frames).

- Frequent frames provide groups of words, but not yet actual syntactic categories;

- Infants still have to figure out which are nouns and which are verbs, possibly using specific concrete words for which they managed to figure out a meaning (e.g; this group contains *bottle*; I know that *bottle* refers to an object; all words from this group also refer to objects)
Can infants use the syntactic category of a new word to constrain its meaning?

<table>
<thead>
<tr>
<th></th>
<th>Video</th>
<th>Audio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Familiarization</td>
<td>Apple turning</td>
<td>Regarde, elle dase! <em>(look, it’s dazzling)</em></td>
</tr>
<tr>
<td>Test</td>
<td>Two apples, one turns, one does something else</td>
<td>Montre-moi celle qui dase! <em>(show me the one that dazzes!)</em></td>
</tr>
</tbody>
</table>

Response: pointing (infants are trained to point beforehand on known words, both objects and actions)
Results: 23-month-old French infants

Show me the one that dazzles!

Savita Bernal (PhD thesis); in collaboration with Jeff Lidz.

Bernal, Lidz, Millotte & Christophe (2007) Syntax constrains the acquisition of verb meaning. *Language Learning and Development*
Control group:

<table>
<thead>
<tr>
<th></th>
<th>Video</th>
<th>Verb group</th>
<th>Noun group (control)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Familiarization</td>
<td>Apple turning</td>
<td>Regarde, elle dase !</td>
<td>Regarde la dase !</td>
</tr>
<tr>
<td></td>
<td>(look, it’s dazzling !)</td>
<td>(look at the dazz !)</td>
<td></td>
</tr>
<tr>
<td>Test</td>
<td>Two apples, one turns, one does something else</td>
<td>Montre-moi celle qui dase ! (show me the one that dazzes !)</td>
<td>Montre-moi la dase ! (show me the dazz !) (stupid question)</td>
</tr>
</tbody>
</table>
Results: 23-month-old French infants (16 in each group)

Savita Bernal (PhD thesis), in collaboration with Jeff Lidz.

The syntactic category of an unknown word constrains its meaning

- When they hear a new word in a variety of constructions (involving only function words and attention-getters), 2-year-olds are able to infer the syntactic category of the new word, and therefore constrain its possible meaning.

- How accurate is syntactic processing at 2 years of age? Debate in literature… task problem
  - Production: don't produce many 'sentences' before age 2;5 - 3 (productive use // imitation?)
  - Comprehension: difficult to interpret looking times.
Evoked potential experiment
relying on article/object pronoun ambiguity
in French

Syntactic category

<table>
<thead>
<tr>
<th>Verb</th>
<th>Noun</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alors elle la mange (Then she eats it)</td>
<td>La poule prend la fraise (The chicken takes the strawberry)</td>
</tr>
</tbody>
</table>

Correct

Incorrect

**La fille prend la mange (The girl takes the eat) **

**Alors il la fraise (Then he strawberry it)**

- No particular task: passive listening
- Use of known words only (from CDI questionnaire)
- To keep infants’ attention focussed, the speaker is playing with toy objects (e.g. strawberry) while she tells a short story; only her face is visible when she utters the test sentences.
Example of Script

Sur ma table, je vois une girafe (N) qui va à l’école. Elle regarde (V) la poule

1. Donc la poule la regarde aussi.  (Correct)

2. Pourtant, elle la girafe très vite!  (Incorrect)

On my table, I see a giraffe (N) who goes to school. She looks (V) at the hen.

1. So the hen looks at her too.  (Correct)

2. However, she giraffes it really fast!  (Incorrect)
Method

- 27 children (mean age 24 months)
- ERP recording: EGI topographic system (128 electrodes)
- Mean good trials/subject: 120
- Collaboration with Ghislaine Dehaene-Lambertz
Infants detect the incorrect sentences

Bernal, Dehaene-Lambertz & Christophe (submitted). Two-year-olds compute syntactic structure on-line *PNAS.*
Distinct neural networks for nouns and verbs

Bernal, Dehaene-Lambertz & Christophe (submitted). Two-year-olds compute syntactic structure on-line *PNAS.*
Distinct neural networks for nouns and verbs

2-year-olds compute syntactic structure on-line.

- 2-year-olds are able to build syntactic structures (they do not simply react to transition probabilities). They build on-line syntactic expectancies: 'je la' predicts a verb, whereas 'je prends la' predicts a noun; Infants are not fooled by the article/object clitic homophony

- Noun and Verb categories have distinct neural bases in 2-year-olds: this is consistent with what we know about adults’ Noun/Verb neural organization (e.g. Damasio & Tranel 1993; Shapiro et Caramazza 2006).
Phonological phrases and function words together:

- Syntax: \([\text{The little boy}]_{NP} [\text{is eating} [\text{an apple}]_{NP}]_{VP}\)
- Prosody: \([\text{The little boy}]_{PP} [\text{is eating}]_{PP} [\text{an apple}]_{PP}\)
- Phrasal prosody and function words together:
  \([\text{the little boy}]_{NP} [\text{is eating}]_{VP} [\text{an apple}]_{NP}\)
  (brackets given by prosody, labels by function words)
- even without content words:
  \([\text{the xxx}]_{NP} [\text{is xx}]_{VP} [\text{an xx}]_{NP}\)
Joint use of phonological phrases and function words: jabberwocky with adults.

- With function word: targets are directly preceded by a function word
  - Noun Sentence: [une **bamoule**] [dri se froliter dagou]
    ([une expo]_NP [doit se dérouler demain])
  - Verb Sentence: [**tu bamoules**] [saman ti] [à mon ada]
    ([tu travailles]_VP [souvent trop] [à mon avis])

- Without function word: targets are not directly preceded by a function word (use of prosodic boundary)
  - Noun Sentence: [sa cramona **bamoule**] [camiche dabou]
    ([sa formidable expo]_NP [commence demain]_VP)
  - Verb Sentence: [sa cramona] [**bamoule** muche]...
    ([sa camarade]_NP [travaille mieux]_VP...)
Abstract word detection task:
ex: ‘bamouler’ (to blick) ‘une bamoule’ (a blick)

Mean percentage of noun and verb responses
to each sentence in each experimental condition

Mean reaction time: 650ms (fast)

The syntactic skeleton

Building expectations as to next constituents

Function word identification and constituent labelling

Phonetic and prosodic analysis

"the little boy runs fast"

"[sa cramona] [bamoule muche]"
Why are noun sentences so difficult:

• Without function word condition – noun sentences:
  [sa cramona bamoule] […]
• Participants respond at chance: 50%
• Two interpretations:
  (1) no prosodic boundary does not mean no syntactic boundary;
  (2) [sa cramona bamoule] can be [det adj noun]
      ex: sa formidable expo (his great exhibition)
      but can also be [det noun adj] (more frequent)
      ex: sa prosodie bizarre (his strange prosody)
Homophony article / object clitic: la pirdale N? V?

Abstract word detection task:
ex: ‘pirdaler’ (to blick)
‘une pirdale’ (a blick)

Elodie Cauvet,
Master’s thesis,
with Séverine Millotte

Je la pirdale_V (I blick_V it) // hier la pirdale_N (yesterday the blick_N…)

Le bamoule_N qui la pirdale_V (the dax_N who blicks_V it) //
Le bamoule_N que la pirdale_N muche (the dax_N whom the blick_N gorps…)

La bamoule_N la pirdale_V digument (the dax_N blicks_V it dithely)
La bamoule_N, la pirdale_N et le taruche muchent (the dax, the blick_N and…)
Bootstrap lexical and syntactic acquisition:

• Joint use of phonological phrases and function words (together with their categories): skeleton of a syntactic structure

  \[ \text{[the xxx]}_{\text{NP}} \text{[is xx]}_{\text{VP}} \text{[an xx]}_{\text{NP}} \]

  maybe around 14-18 months...

  \[ \Rightarrow \text{potentially enough information to constrain} \]

  the acquisition of word meanings.
Model of processing and acquisition

Syntactic processing

[the xxx]_{NP} [is xx]_{VP} [an xx]_{NP}

Phonological phrases and function words

Function words

Phonological and prosodic representation

Speech signal

"the little boy is eating an apple"
Bootstrap lexical and syntactic acquisition:

• Triple research strategy:
  - corpus studies (feasibility);
  - experiments with adults (on-line studies);
  - experiments with infants (acquisition);

• Cross-linguistic studies.
Joint use of phonological phrases and function words: jabberwocky with adults.

- Nonword targets are directly preceded by a function word, article or pronoun:
  - Noun Sentence: [une bamoule] [dri se froliter dagou]
    ([une expo]_{NP} [doit se dérouler demain])
  - Verb Sentence: [tu bamoules] [saman ti] [à mon ada]
    ([tu travailles]_{VP} [souvent trop] [à mon avis])

- Without function word: targets are not directly preceded by a function word (use of prosodic boundary)
  - Noun Sentence: [sa cramona bamoule] [camiche dabou]
    ([sa formidable expo]_{NP} [commence demain]_{VP})
  - Verb Sentence: [sa cramona] [bamoule muche]…
    ([sa camarade]_{NP} [travaille mieux]_{VP}…)
Joint use of phonological phrases and function words: jabberwocky with adults.

- Nonword targets are directly preceded by a function word, article or pronoun: target 'bamoule'
  - Noun Sentence: [une bamoule] [dri se froliter dagou]
    ([une expo]_{NP} [doit se dérouler demain])
    (an exposition will take place tomorrow)
  - Verb Sentence: [tu bamoules] [saman ti] [à mon ada]
    ([tu travailles]_{VP} [souvent trop] [à mon avis])
    (you often work too much in my opinion)

- Task: abstract word detection: ex: 'bamouler' to blick
  'une bamoule' a blick

Verb target/Verb sentence: press a button; Verb target/Noun sentence: refrain from responding.
Abstract word detection task: ex: ‘bamouler’ (to blick)  ‘une bamoule’ (a blick)

Mean percentage of noun and verb responses to each sentence in each experimental condition

Mean reaction time: 650ms (fast)

Are prosodic cues spontaneously produced by naïve speakers? Yes

The same pattern obtained for every single speaker

Millotte, René, Wales & Christophe (in revision) *Language & Cognitive Processes*