

# Je jazyk vrožený?

## Debata mezi nativismem a empiricismem

---

LGV22 PSYCHOLINGVISTIKA  
PŘEDNÁŠKA 2

# Opakování z minula

---

Moderní lingvistika vnímá jazyk jako **individuální mentální schopnost**

Lingvisté zkoumají **gramatiku** jazyků

- Implicitní **znalost pravidel**, které umožňují generaci nekonečného množství možných vět

Kognitivní systémy lze zkoumat na několika úrovních (Marr, 1982)

- Formální lingvisté (sémantici, syntaktici, fonologové) zkoumají jazyk na **komputační úrovni**

Psycholingvistika se zabývá **algoritmickou úrovní**

- Zkoumáme **procesy a reprezentace** vyskytující se v osvojení a produkci jazyka a v jeho porozumění.

# Dnešní hodina

---

1. Problém osvojení jazyka a jeho význam pro lingvistickou teorii
2. Nativismus vs. empiricismus
3. Argument z chudoby stimulu
4. Zjišťovací otázky v angličtině
5. Statistické učení a jeho omezení

# Osvojení jazyka



# Osvojení jazyka (akvizice)

---

Je proces, který umožňuje jednotlivci získat buď svůj mateřský (L1) nebo druhý jazyk (L2).



# Osvojení jazyka (akvizice)

---

Je proces, který umožňuje jednotlivci získat buď svůj mateřský (L1) nebo druhý jazyk (L2).

V této a příští hodině se budeme soustředit na **osvojení mateřského jazyka** (first language acquisition).



# Osvojení jazyka (akvizice)

---

Je proces, který umožňuje jednotlivci získat buď svůj mateřský (L1) nebo druhý jazyk (L2).

V této a příští hodině se budeme soustředit na **osvojení mateřského jazyka** (first language acquisition).

Tento týden se podíváme na otázku **jazykového nativismu**.



# Proč je zkoumání akvizice důležité?

---

Proč si všechny zdravě se vyvíjející děti dokáží osvojit jazyk v tak krátké době a bez zvláště velkého úsilí?

Proč si lidé, kteří nebyli jazyku vystaveni v dětství, jej nedokáží v dospělosti osvojit?

Proč se žádné zvíře nedokáže naučit lidský jazyk i přes značný trénink?



# Proč je zkoumání akvizice důležité?

---

Proč si všechny zdravě se vyvíjející děti dokáží osvojit jazyk v tak krátké době a bez zvláště velkého úsilí?

Proč si lidé, kteří nebyli jazyku vystaveni v dětství, jej nedokáží v dospělosti osvojit?

Proč se žádné zvíře nedokáže naučit lidský jazyk i přes značný trénink?

**Zkoumání osvojení jazyka nám pomáhá zjistit, co to vlastně jazyk je!**

# Nativismus



# Nativismus

---

Běžný význam slova *vrozený* by se dal zachytit například těmito definicemi (Griffiths, 2017):

- To, s čím se člověk narodí
- To, co je v genech

Pokud je psychologická vlastnost vrozená, pak **není naučená** (Samuels, 2002)

# Nativismus

---

Běžný význam slova *vrozený* by se dal zachytit například těmito definicemi (Griffiths, 2017):

- To, s čím se člověk narodí
- To, co je v genech

Pokud je psychologická vlastnost vrozená, pak **není naučená** (Samuels, 2002)

V kontextu jazyka je tedy nativismus pozice, která tvrdí, že existuje určitá **jazyková znalost**, která je vrozená a ne naučená.

**Empiricismem** pak budeme mínit pozici, která tvrdí, že tomu tak není.

# Berwick et al. (2011)

---

Čtyři možné faktory:

1. Vrozené, doménově specifické faktory
2. Vrozené, doménově všeobecné faktory
3. Externí stimuly
4. Přírodní zákony

# Berwick et al. (2011)

---

Čtyři možné faktory:

1. Vrozené, **doménově specifické** faktory
2. Vrozené, doménově všeobecné faktory
3. Externí stimuly
4. Přírodní zákony

# Berwick et al. (2011)

---

Čtyři možné faktory:

1. Vrozené, doménově specifické faktory
2. Vrozené, **doménově všeobecné** faktory
3. Externí stimuly
4. Přírodní zákony

# Argument z chudoby stimulu

---

Argument ve prospěch nativismu v rámci osvojení jazyka



# Argument z chudoby stimulu

---

Argument ve prospěch nativismu v rámci osvojení jazyka

Jako u každého jiného argumentu se musíme ptát:

- Je **validní**? → logická „správnost“
- Jsou jeho premisy pravdivé?

# Argument z chudoby stimulu

---

Argument ve prospěch nativismu v rámci osvojení jazyka

Jako u každého jiného argumentu se musíme ptát:

- Je **validní**? → logická „správnost“
- Jsou jeho premisy pravdivé?

Děti vykazují jazykové znalosti, které se **nemohly naučit**. Z toho vyplývá, že tyto znalosti musí být vrozené.

# Analýza od Laurence a Margolise (2001)

---

1. Neurčité množství alternativních množin principů je konsistentní s pravidelnostmi, které se nachází v primárních jazykových datech.

# Analýza od Laurence a Margolise (2001)

---

1. Neurčité množství alternativních množin principů je konsistentní s pravidelnostmi, které se nachází v primárních jazykových datech.
2. Ta správná množina principů nemusí být (a typicky není) nikterak jednodušší nebo přirozenější v preteoretickém smyslu než další alternativy.

# Analýza od Laurence a Margolise (2001)

---

1. Neurčité množství alternativních množin principů je konsistentní s pravidelnostmi, které se nachází v primárních jazykových datech.
2. Ta správná množina principů nemusí být (a typicky není) nikterak jednodušší nebo přirozenější v preteoretickém smyslu než další alternativy.
3. Data, která by byla potřeba pro vybrání té správné množiny principů, nejsou v mnoha případech k dispozici empiristickému osvojovateli („empiricist learner“) v epistemické situaci dítěte.

# Analýza od Laurence a Margolise (2001)

---

1. Neurčité množství alternativních množin principů je konsistentní s pravidelnostmi, které se nachází v primárních jazykových datech.
2. Ta správná množina principů nemusí být (a typicky není) nikterak jednodušší nebo přirozenější v preteoretickém smyslu než další alternativy.
3. Data, která by byla potřeba pro vybrání té správné množiny principů, nejsou v mnoha případech k dispozici empiristickému osvojovateli („empiricist learner“) v epistemické situaci dítěte.
4. Pokud by děti byly empiristickými osvojovateli, tak si by nemohli spolehlivě osvojit tu správnou gramatiku pro svůj jazyk.

# Analýza od Laurence a Margolise (2001)

---

1. Neurčité množství alternativních množin principů je konsistentní s pravidelnostmi, které se nachází v primárních jazykových datech.
2. Ta správná množina principů nemusí být (a typicky není) nikterak jednodušší nebo přirozenější v preteoretickém smyslu než další alternativy.
3. Data, která by byla potřeba pro vybrání té správné množiny principů, nejsou v mnoha případech k dispozici empiristickému osvojovateli („empiricist learner“) v epistemické situaci dítěte.
4. Pokud by děti byly empiristickými osvojovateli, tak si by nemohli spolehlivě osvojit tu správnou gramatiku pro svůj jazyk.
5. Děti si dokáží spolehlivě osvojit gramatiku svého jazyka.

# Analýza od Laurence a Margolise (2001)

---

1. Neurčité množství alternativních množin principů je konsistentní s pravidelnostmi, které se nachází v primárních jazykových datech.
2. Ta správná množina principů nemusí být (a typicky není) nikterak jednodušší nebo přirozenější v preteoretickém smyslu než další alternativy.
3. Data, která by byla potřeba pro vybrání té správné množiny principů, nejsou v mnoha případech k dispozici empiristickému osvojovateli („empiricist learner“) v epistemické situaci dítěte.
4. Pokud by děti byly empiristickými osvojovateli, tak si by nemohli spolehlivě osvojit tu správnou gramatiku pro svůj jazyk.
5. Děti si dokáží spolehlivě osvojit gramatiku svého jazyka.
6. Tudíž děti nejsou empiristickými osvojovateli.



# Zjišťovací otázky v angličtině





The cat is on the roof.

---



?

The cat is on the roof.

---



The cat who is white is on the roof.

---



The cat who is white is on the roof.





\*Is the cat who white is on the roof?

---



Ale děti osvojující si angličtinu **nikdy**  
podobné chyby nedělají!  
(Crain & Nakayama, 1987)

\*Is the cat who white is on the roof?

---

# Jak vypadá struktura vět?

---

The cat is on the roof.



# Jak vypadá struktura vět?

---

The cat is on the roof.

1 2 3 4 5 6

# Jak vypadá struktura vět?

---

*Lineární struktura*

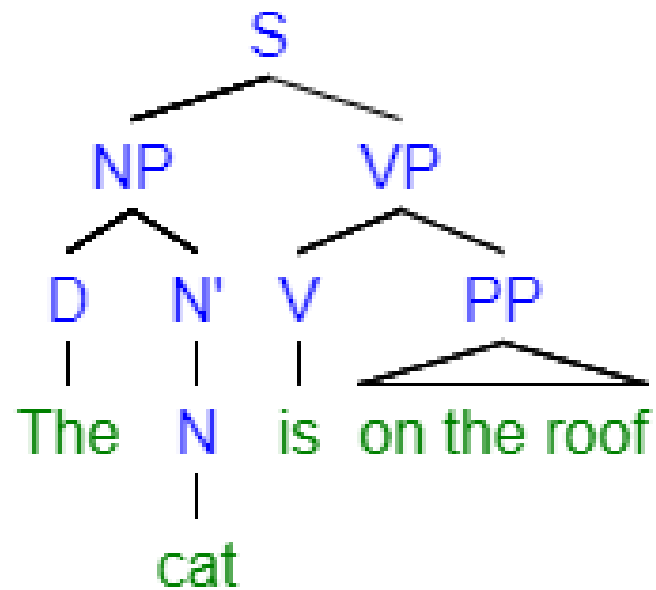
The cat is on the roof.

1 2 3 4 5 6

# Jak vypadá struktura vět?

---

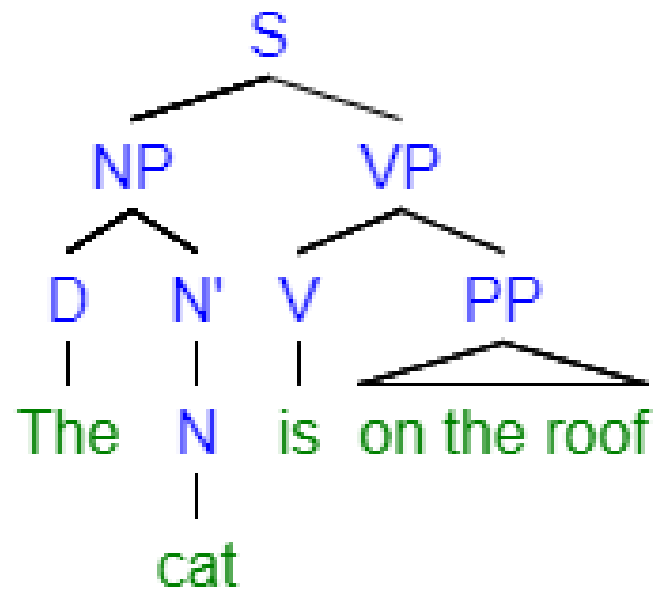
## *Hierarchická struktura*



# Jak vypadá struktura vět?

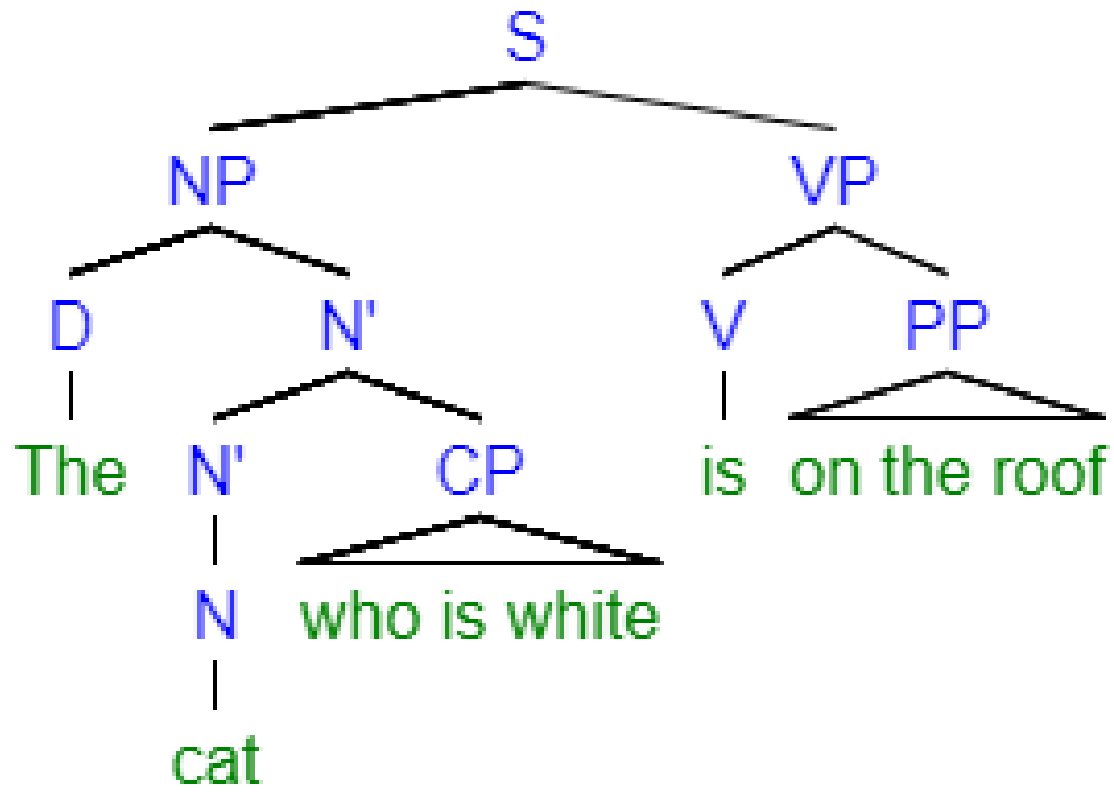
---

## *Hierarchická struktura*



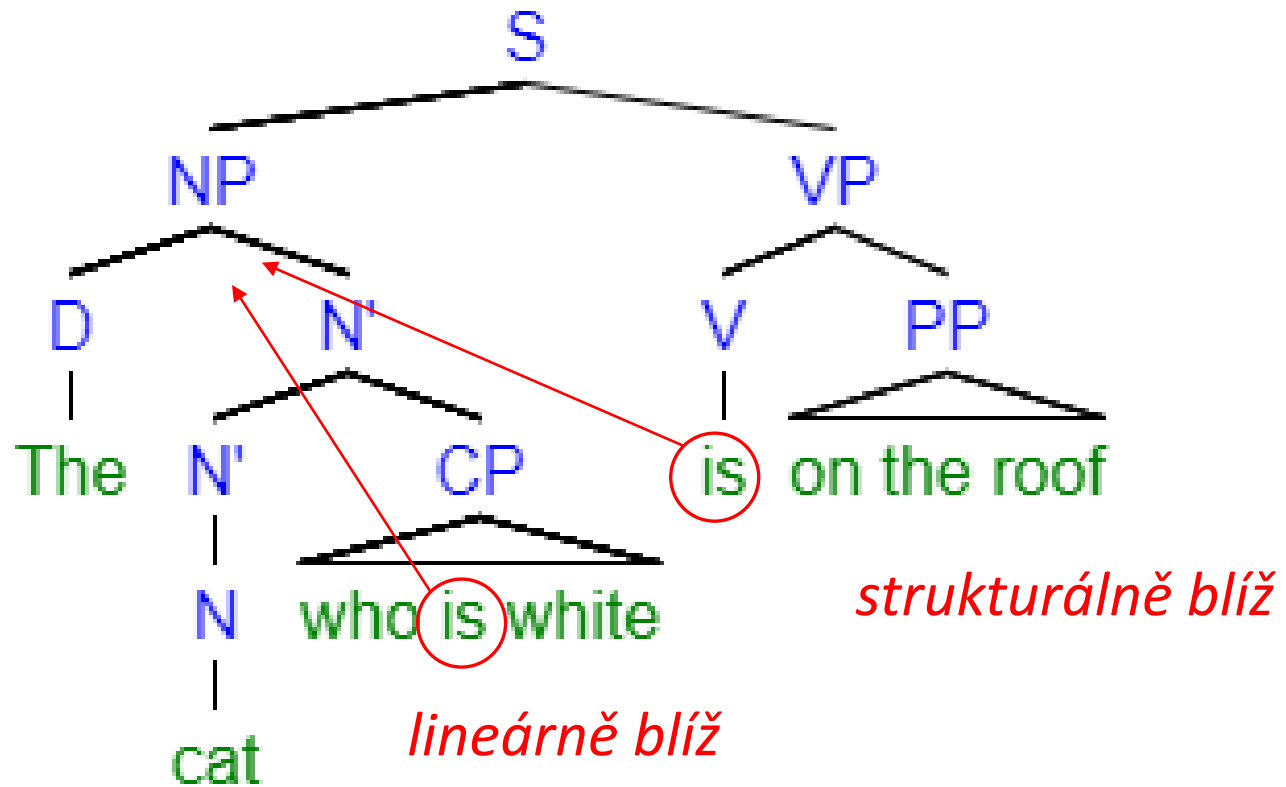
# Hierarchická struktura

---



# Hierarchická struktura

---



Ta zásadní otázka je...

---

# Ta zásadní otázka je...

---

Mají děti dostatek dat k tomu, aby se naučily, že větná struktura je *hierarchická* a ne *lineární*?



# Ta zásadní otázka je...

---

Mají děti dostatek dat k tomu, aby se naučily, že větná struktura je *hierarchická* a ne *lineární*?

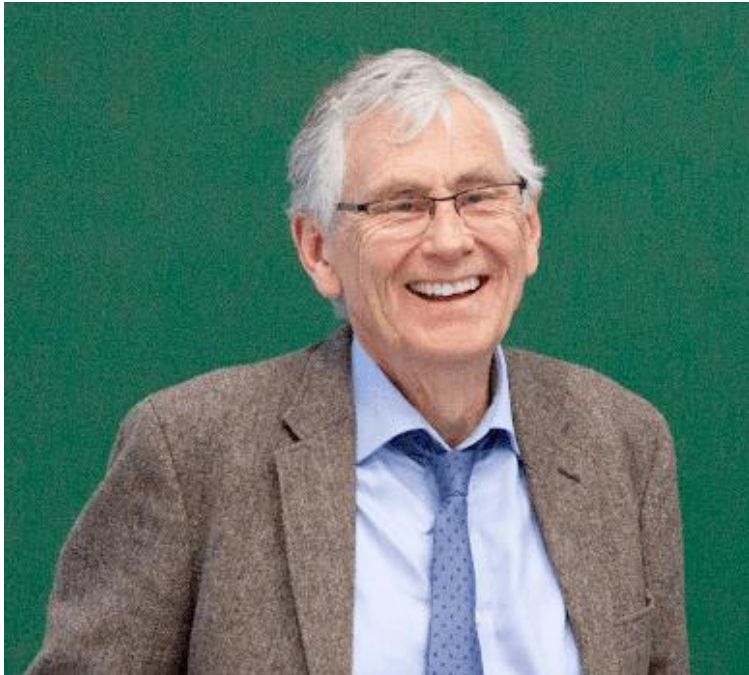
*“A person might go through much or all of his life without ever having been exposed [to such sentences]”* – Noam Chomsky (see Piattelli-Palmarini 1980: 40)

# Pullum & Scholz (2002)

---

## Empirical assessment of stimulus poverty arguments<sup>1</sup>

GEOFFREY K. PULLUM AND BARBARA C. SCHOLZ



### *Abstract*

*This article examines a type of argument for linguistic nativism that takes the following form: (i) a fact about some natural language is exhibited that allegedly could not be learned from experience without access to a certain kind of (positive) data; (ii) it is claimed that data of the type in question are not found in normal linguistic experience; hence (iii) it is concluded that people cannot be learning the language from mere exposure to language use. We analyze the components of this sort of argument carefully, and examine four exemplars, none of which hold up. We conclude that linguists have some additional work to do if they wish to sustain their claims about having provided support for linguistic nativism, and we offer some reasons for thinking that the relevant kind of future work on this issue is likely to further undermine the linguistic nativist position.*

---

1. This work was partially supported by grants from the Humanities Division at the University of California, Santa Cruz. Some of the material was presented in an earlier form by Pullum in the Parasession on Learnability at the Berkeley Linguistics Society's annual meeting in February 1996, and appeared as Pullum (1996). Early versions of some of the research



# Reali & Christiansen (2005)

---

Cognitive Science 29 (2005) 1007–1028  
Copyright © 2005 Cognitive Science Society, Inc. All rights reserved.



## Uncovering the Richness of the Stimulus: Structure Dependence and Indirect Statistical Evidence

Florencia Reali, Morten H. Christiansen

*Department of Psychology, Cornell University*

Received 12 August 2004; received in revised form 9 March 2005; accepted 22 March 2005

---

### Abstract

The *poverty of stimulus argument* is one of the most controversial arguments in the study of language acquisition. Here we follow previous approaches challenging the assumption of impoverished primary linguistic data, focusing on the specific problem of auxiliary (AUX) fronting in complex polar interrogatives. We develop a series of corpus analyses of child-directed speech showing that there is indirect statistical information useful for correct auxiliary fronting in polar interrogatives and that such information is sufficient for distinguishing between grammatical and ungrammatical generalizations, even in the absence of direct evidence. We further show that there are simple learning devices, such as neural networks, capable of exploiting such statistical cues, producing a bias toward correct AUX questions when compared to their ungrammatical counterparts. The results suggest that the basic assumptions of the poverty of stimulus argument may need to be reappraised.

*Keywords:* Poverty of stimulus; Distributional information; Corpus analysis; Neural networks; Language acquisition

---



# Reali & Christiansen (2005)

---

Dokáží děti využít **nepřímých statistických indicií** k tomu, aby si osvojily syntax zjišťovacích otázek v angličtině?

# Reali & Christiansen (2005)

---

Dokáží děti využít **nepřímých statistických indicií** k tomu, aby si osvojily syntax zjišťovacích otázek v angličtině?

Využili **korpus tzv. „child-directed speech“** (Bernstein-Ratner, 1984) k trénování statistických modelů

# Reali & Christiansen (2005)

---

Dokáží děti využít **nepřímých statistických indicií** k tomu, aby si osvojily syntax zjišťovacích otázek v angličtině?

Využili **korpus tzv. „child-directed speech“** (Bernstein-Ratner, 1984) k trénování statistických modelů

Experiment 1 a 2: **bigramový** a **trigramový** model

Experiment 3: **Simple recurrent network** model (SRN)

# Reali & Christiansen (2005)

---

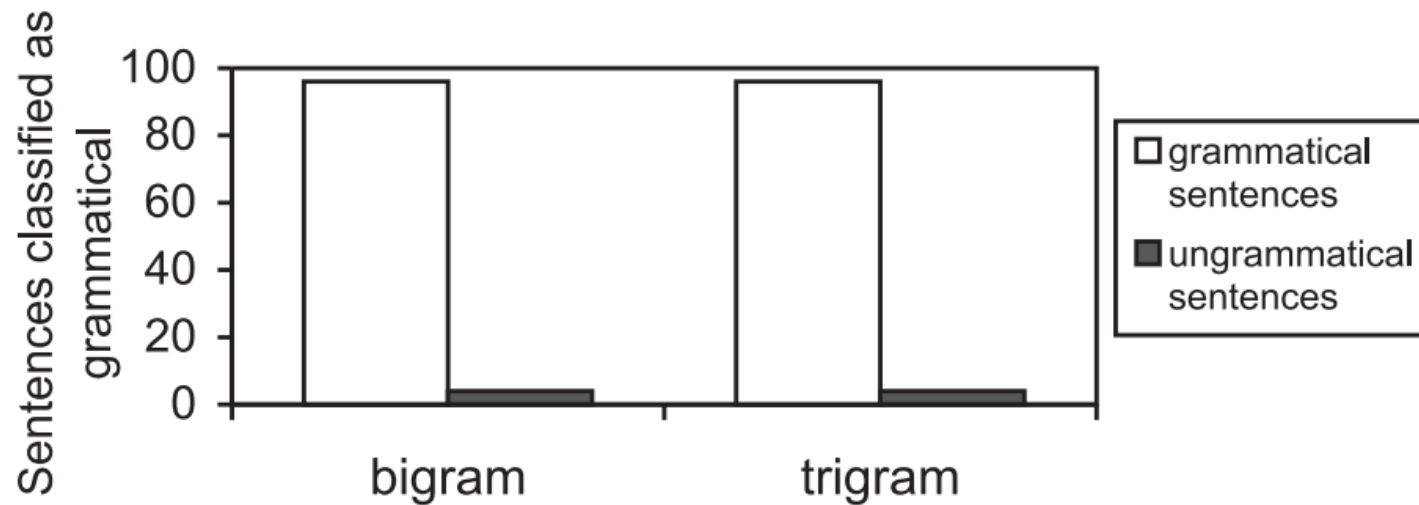


Fig. 2. Number of sentences classified correctly (white bars) and incorrectly as grammatical (black bars).

# Reali & Christiansen (2005)

---

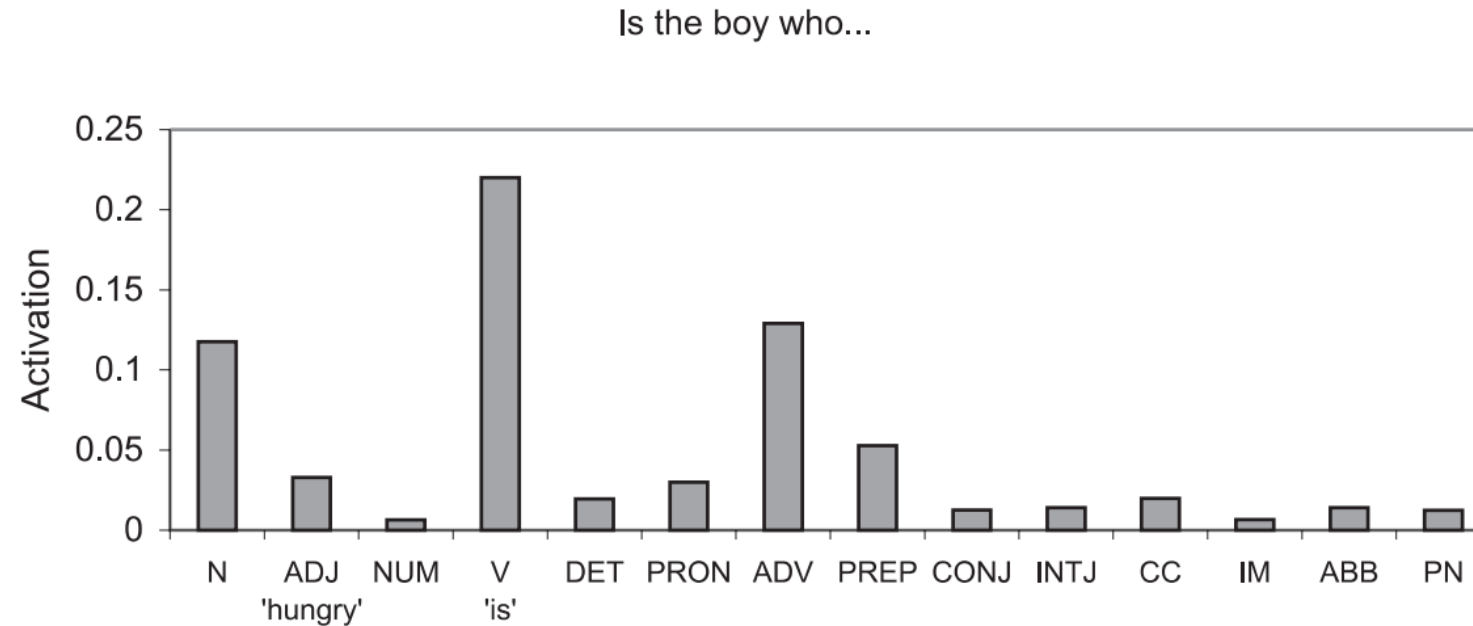


Fig. 5. Network prediction after the presentation of the lexical classes corresponding to: “*Is he boy who ...*” The prediction of the well-formed relative clause continuation, V (i.e., *is*), is highly preferred over the ill-formed version, ADJ (i.e., *hungry*).



# Reali & Christiansen (2005)

---

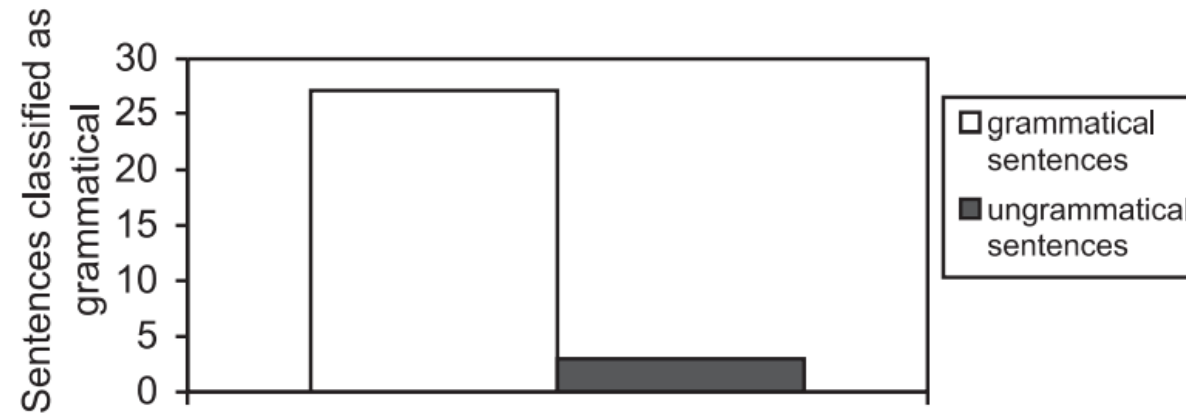


Fig. 6. Number of sentences classified correctly by the SRNs (white bars) and incorrectly as grammatical (black bars).

# Statistické učení u dětí



# Saffran, Aslin, & Newport (1996)

---

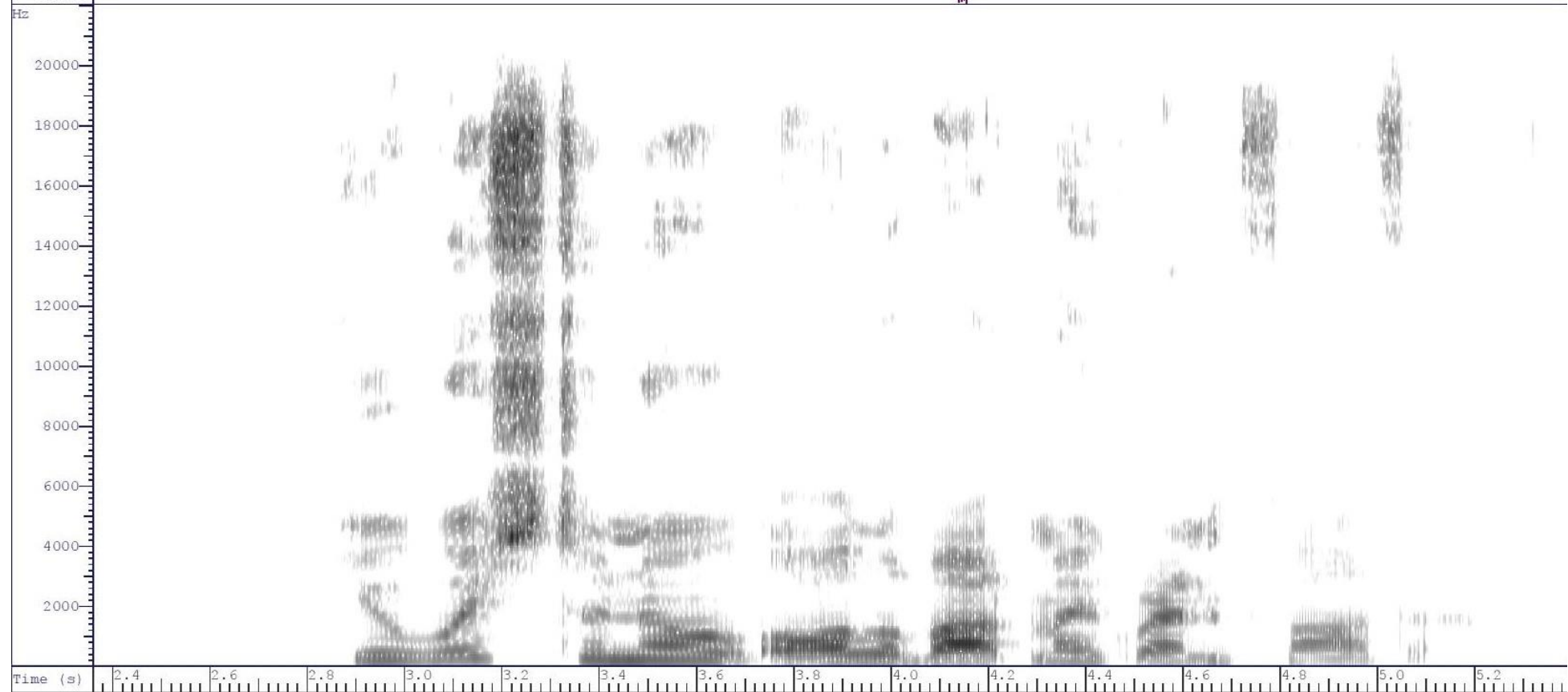
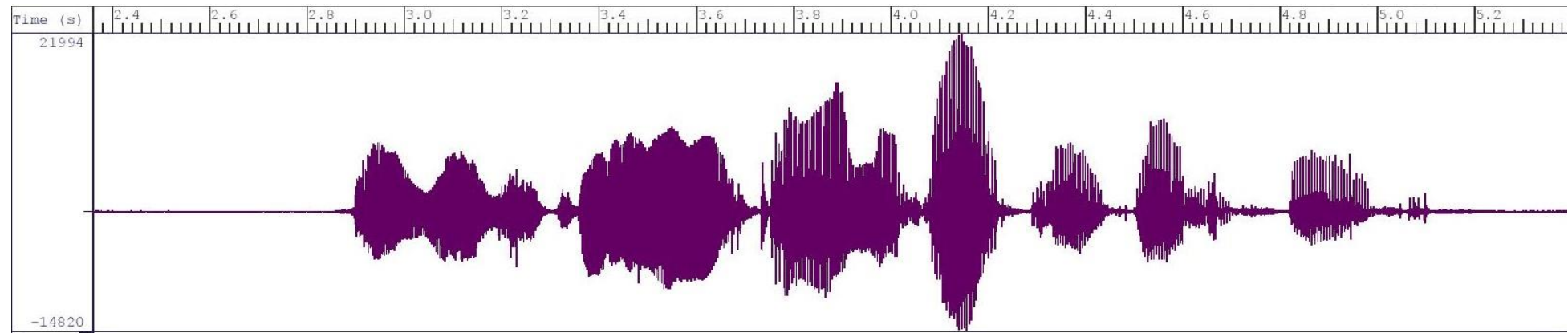


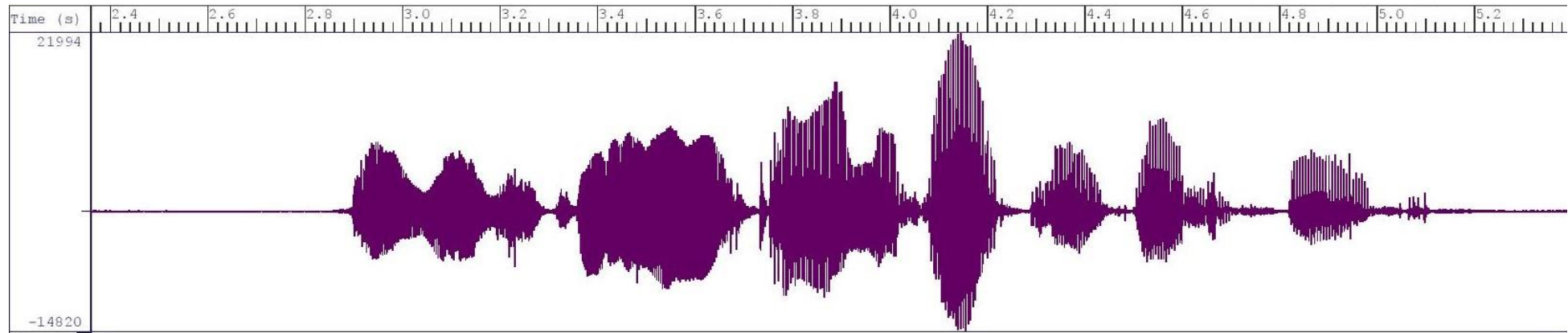
## **Statistical Learning by 8-Month-Old Infants**

Jenny R. Saffran, Richard N. Aslin, Elissa L. Newport

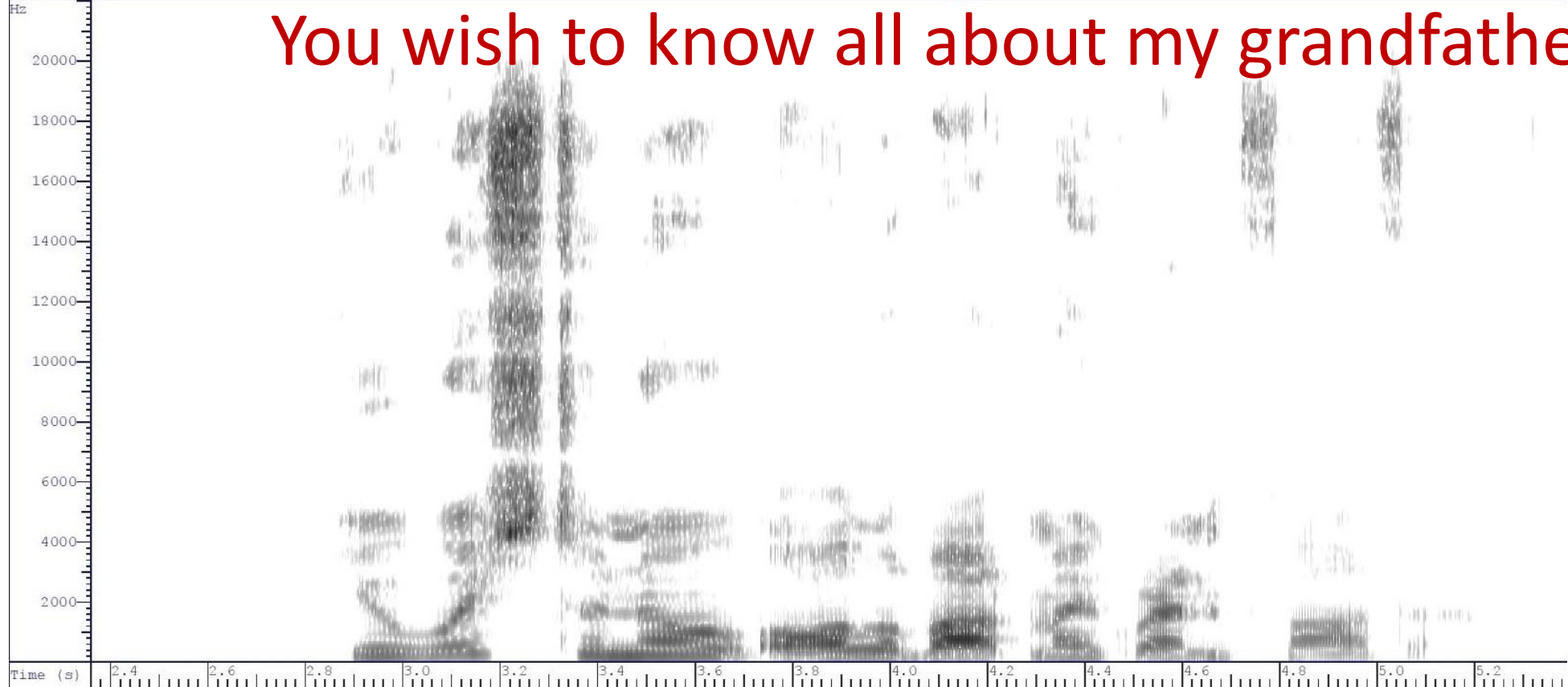
Learners rely on a combination of experience-independent and experience-dependent mechanisms to extract information from the environment. Language acquisition involves both types of mechanisms, but most theorists emphasize the relative importance of experience-independent mechanisms. The present study shows that a fundamental task of language acquisition, segmentation of words from fluent speech, can be accomplished by 8-month-old infants based solely on the statistical relationships between neighboring speech sounds. Moreover, this word segmentation was based on statistical learning from only 2 minutes of exposure, suggesting that infants have access to a powerful mechanism for the computation of statistical properties of the language input.

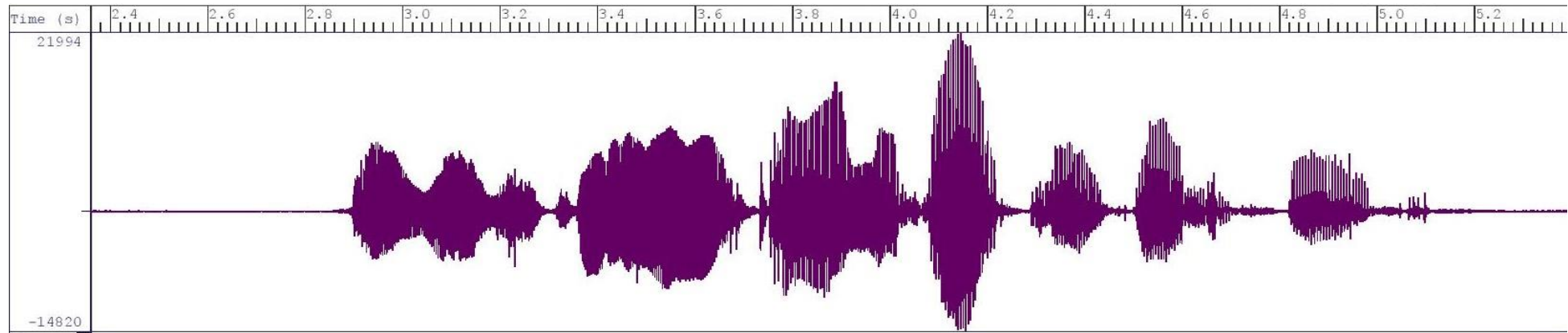
---



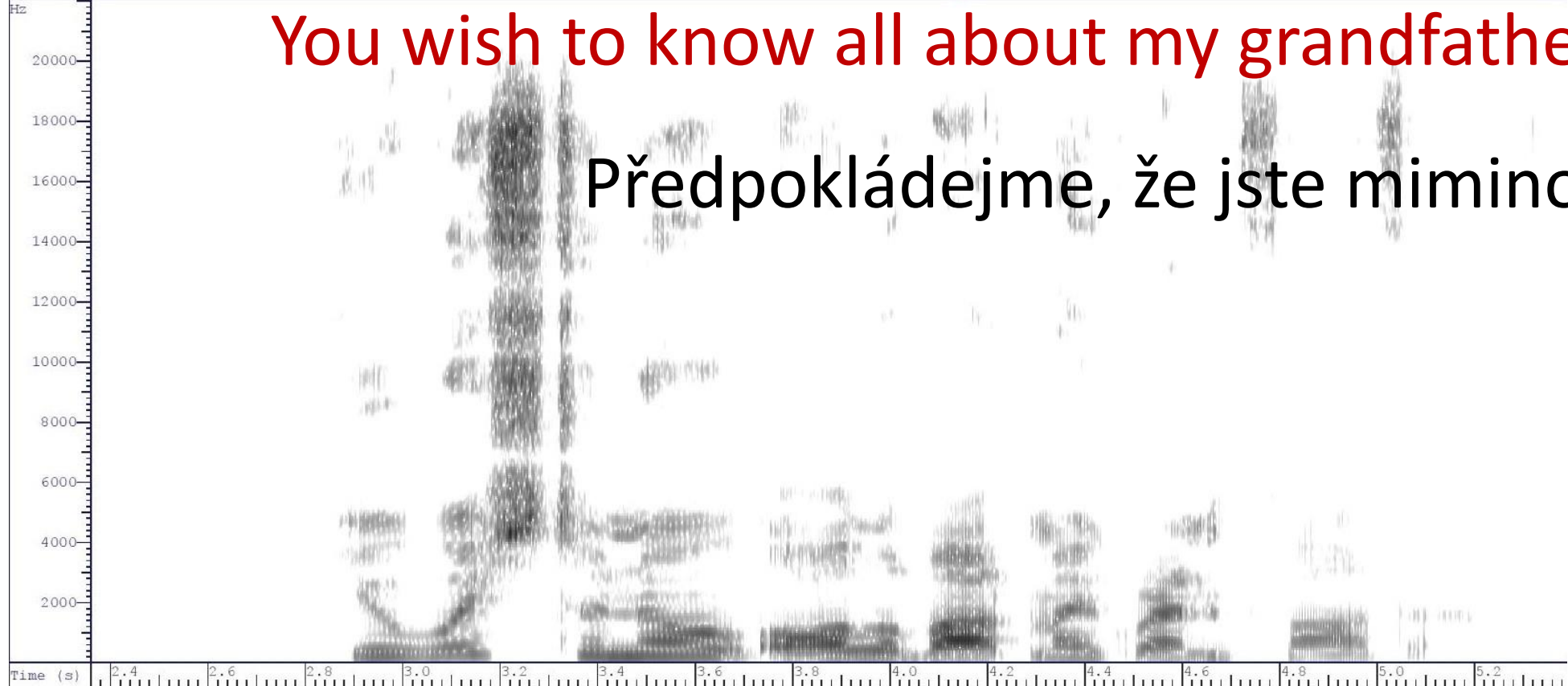


You wish to know all about my grandfather





You wish to know all about my grandfather  
Předpokládejme, že jste mimino...





# Saffran, Aslin, & Newport (1996)

---

Dětem ve věku 8 měsíců byly prezentovány **řetězce slabik** po dobu dvou minut:



# Saffran, Aslin, & Newport (1996)

---

Dětem ve věku 8 měsíců byly prezentovány **řetězce slabik** po dobu dvou minut:

*bidakupadotigolabubidaku*

# Saffran, Aslin, & Newport (1996)

---

Dětem ve věku 8 měsíců byly prezentovány **řetězce slabik** po dobu dvou minut:

***bidakupadotigolabubidaku***



**p = 1.0**

# Saffran, Aslin, & Newport (1996)

---

Dětem ve věku 8 měsíců byly prezentovány řetězce slabik po dobu dvou minut:

$p = 0.33$



*bidakupadotigolabubidaku*



$p = 1.0$

# Saffran, Aslin, & Newport (1996)

---

Dětem ve věku 8 měsíců byly prezentovány **řetězce slabik** po dobu dvou minut

Familiarisation-preference procedure

- Po familiarizaci jsou prezentovány buď **známé** nebo **nové** podněty

Během testu kojenci kontrolují trvání stimulu tím, jak dlouho fixují svůj pohled na blikající světlo

Preferují kojenci **známá slova** nebo **nová neslova**?



# Saffran, Aslin, & Newport (1996)

---

Dětem ve věku 8 měsíců byly prezentovány **řetězce slabik** po dobu dvou minut

Familiarisation-preference procedure

- Po familiarizaci jsou prezentovány buď **známé** nebo **nové** podněty

Během testu kojenci kontrolují trvání stimulu tím, jak dlouho fixují svůj pohled na blikající světlo

Preferují kojenci **známá slova** (*bidaku*) nebo **nová neslova** (*kupada*)?

**Statisticky signifikantně delší doba poslechu pro nová neslova (t.j. „novelty preference“)**

→ Děti dokáží rozlišit mezi těmito dvěma typy

→ Osmiměsíční kojenci dokáží extrahovat a využít statistické informace ze stimulu, i když jsou mu vystaveni velmi krátkou dobu

# Shrnutí

---

1. Zkoumání osvojení jazyka je zásadní pro naši celkovou teorii o jazyce.
2. Nativisté tvrdí, že existuje vrozená znalost jazyka, která je doménově specifická.
3. Argument z chudoby stimulu podporuje nativistickou pozici.
4. Tento argument stojí na tvrzení, že data, která jsou dětem k dispozici, jsou nedostatečná.
5. Ale výzkum v posledních desetiletích ukázal, že děti jsou schopny velmi účtyhodné úrovně statistického učení.

---

Nyní je prostor pro otázky

# Diskuze: Lidz, Waxman & Freedman (2003)

---



Brief article

What infants know about syntax but couldn't have learned: experimental evidence for syntactic structure at 18 months

Jeffrey Lidz<sup>a,\*</sup>, Sandra Waxman<sup>b</sup>, Jennifer Freedman<sup>a</sup>

<sup>a</sup>Department of Linguistics, Northwestern University, 2016 Sheridan Road, Evanston, IL 60208, USA

<sup>b</sup>Department of Psychology, Northwestern University, 2029 Sheridan Road, Evanston, IL 60208, USA

Received 14 November 2002; revised 16 April 2003; accepted 25 May 2003

---

## Abstract

Generative linguistic theory stands on the hypothesis that grammar cannot be acquired solely on the basis of an analysis of the input, but depends, in addition, on innate structure within the learner to guide the process of acquisition. This hypothesis derives from a logical argument, however, and its consequences have never been examined experimentally with infant learners. Challenges to this hypothesis, claiming that an analysis of the input is indeed sufficient to explain grammatical acquisition, have recently gained attention. We demonstrate with novel experimentation the insufficiency of this countervailing view. Focusing on the syntactic structures required to determine the antecedent for the pronoun *one*, we demonstrate that the input to children does not contain sufficient information to support unaided learning. Nonetheless, we show that 18-month-old infants do have command of the syntax of *one*. Because this syntactic knowledge could not have been gleaned exclusively from the input, infants' mastery of this aspect of syntax constitutes evidence for the contribution of innate structure within the learner in acquiring a grammar.

© 2003 Elsevier B.V. All rights reserved.

*Keywords:* Universal grammar; Poverty of the stimulus; Language acquisition

---

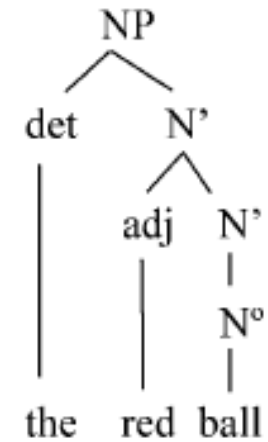
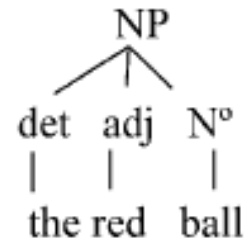


# Diskuze: Lidz, Waxman & Freedman (2003)

---

(1)

a. Flat structure hypothesis    b. Nested structure hypothesis



# Diskuze: Lidz, Waxman & Freedman (2003)

---

*J. Lidz et al. / Cognition 89 (2003) B65–B73*

B71

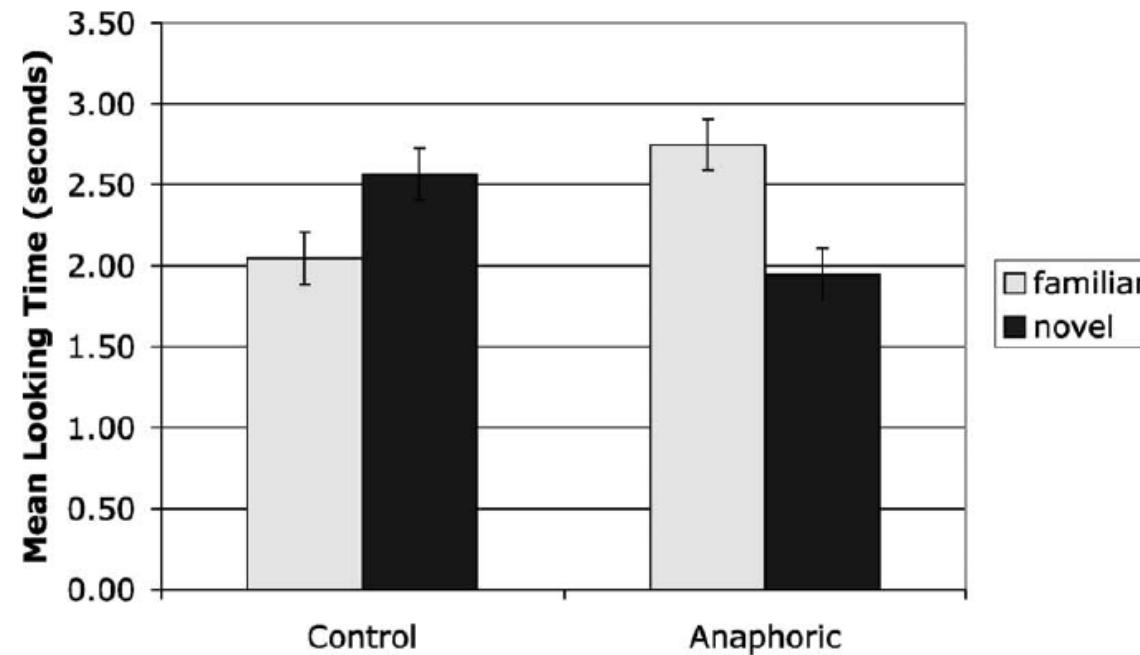


Fig. 1. Mean looking time (in seconds) to the two test images in each condition.