The Syntax of Birdsong

Berwick, Okanoya, Beckers & Bolhuis

2011

Zebra finch song: hierarchical structure

recording

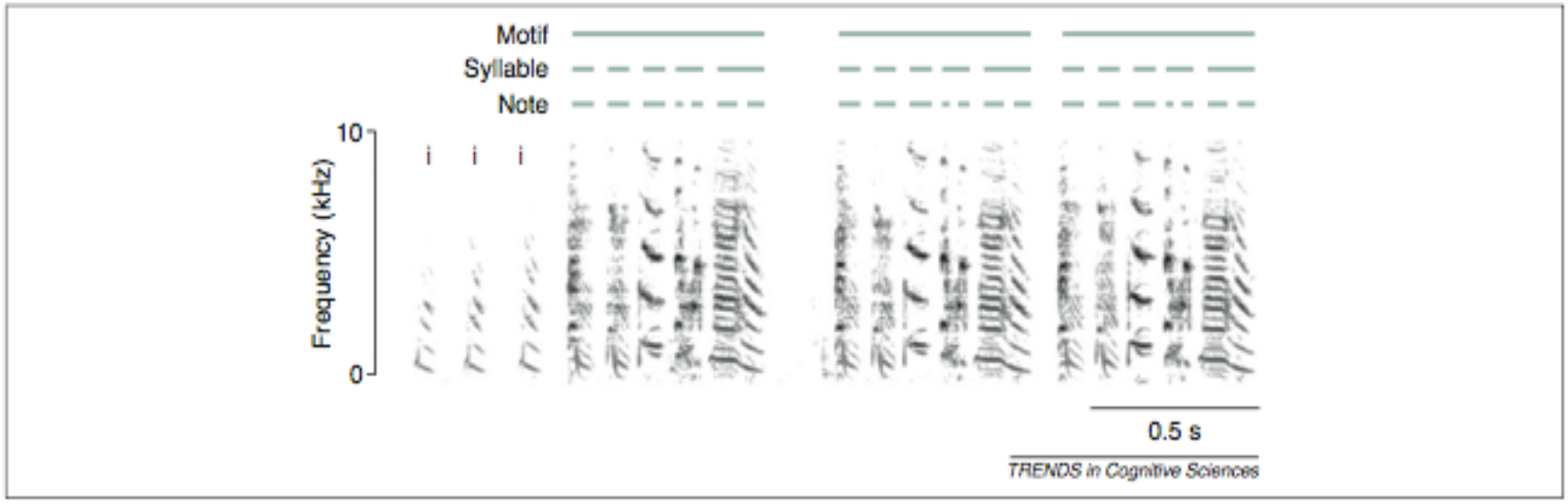


Figure 1. Sound spectrogram of a typical zebra finch song depicting a hierarchical structure. Songs often start with 'introductory notes' (denoted by 'i') that are followed by one or more 'motifs', which are repeated sequences of syllables. A 'syllable' is an uninterrupted sound, which consists of one or more coherent time-frequency traces, which are called 'notes'. A continuous rendition of several motifs is referred to as a 'song bout'.

Structure of birdsong

- Birdsong consists of chains of discrete acoustic elements arranged in a particular temporal order.
- Songs might consist of fixed sequences with only sporadic variation (e.g. zebra finches), or more variable sequences.
- A song of a nightingale is built out of a fixed 4-second note sequence.
  - An individual nightingale has 100-200 song types, clustered into 2-12 ‘packages’.
  - Recording.

Do bird songs have phonology or syntax?

- Berwick et al. definitions:
  - Phonetics:
    - The study of the actual speech sounds of all languages, including their physical properties, the way they are perceived and the way in which vocal organs produce sounds.
  - Phonology
    - The study of the abstract sound patterns of a particular language, usually according to some system of rules.
  - Syntax
    - The rules for arranging items (sounds, words, word parts or phrases) into their possible permissible combinations in a language.
  - Semantics
    - the analysis of the meaning of a language at the word, phrase, sentence level, or beyond.

Birdsong only exhibits “phonological syntax”

- Since bird songs do not involve combinations of sound elements into morphemes, morphemes into words, and words into phrases and sentences, they should be characterized as only exhibiting “phonological syntax”.
  - This is because bird song lacks semantics: song elements are not combined to form novel ‘meanings’.
  - Instead bird song can only convey a limited set of intentions, as a graded, holistic communication system to attract mates or deter rivals and defend territory.
- We can still investigate the combinatorial system of bird song and discover how complex it is compared to human language phonology and syntax.

Structure of a generative grammar: Chomsky 1995: *The Minimalist Program*

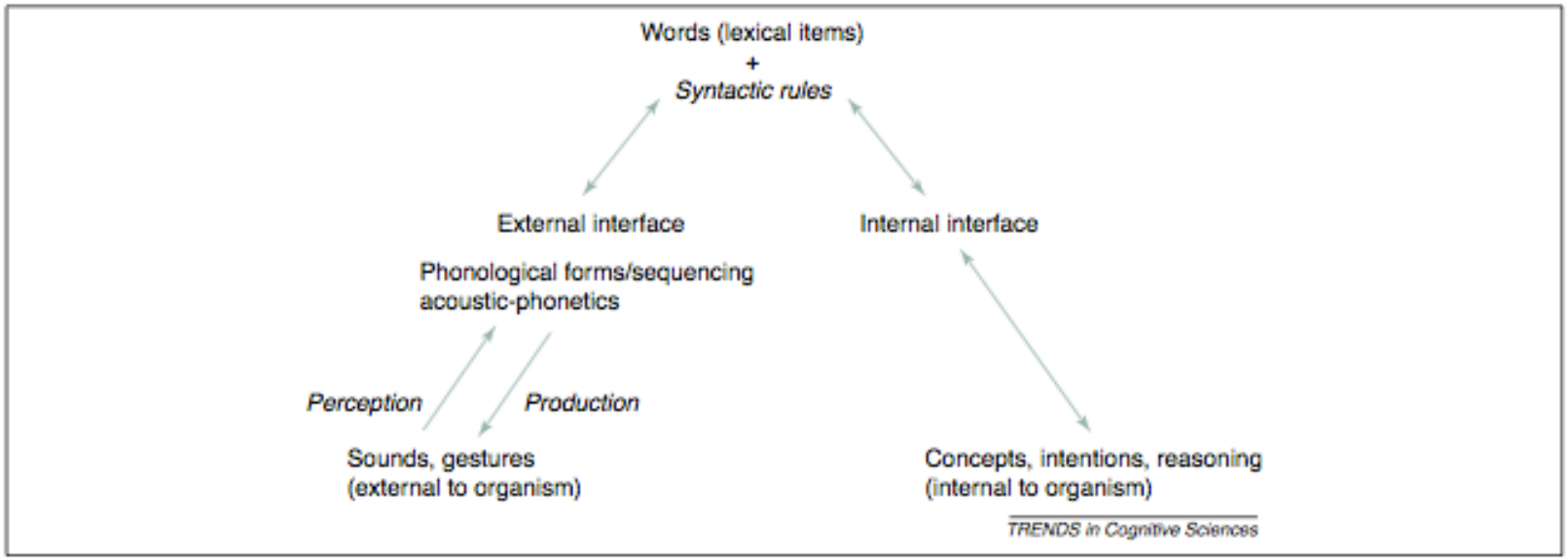


Figure 2. A tripartite diagram of abstract components encompassing both human language and birdsong. On the left-hand side, an external interface **(i)**, comprised of sensorimotor systems, links the perception and production of acoustic signals to an internal system of syntactic rules, **(ii)**. On the right-hand side, an internal interface links syntactic forms to some system of concepts and intentions, **(iii)**. With respect to this decomposition, birdsong seems distinct from human language in the sense of lacking both words and a fully developed conceptual–intentional system.

Chomsky 1957: *Syntactic Structures*

- Proposed the idea of a generative grammar.
  - A language is an (infinite) set of well-formed sentences (sequences of words).
  - The goal is to discover a fully explicit, formal mathematical device that can generate all and only the well-formed strings of a given language.
- Stepped through a series of formal grammars with different types of rules.
  - Showed that some aspects of English can be accounted for by rules characteristic of the simplest type of grammar: a finite state grammar.
  - Showed that some aspects of English require a more complex type of rule: context-free phrase structure rules.
  - Argued that still more complex, context-sensitive phrase structure rules might be needed for some aspects of English, and that still more complex rules (transformational rules) are needed to insightfully describe English.

Legacy of *Syntactic Structures*

- Chomsky's idea that there is a hierarchy of formal grammars and corresponding formal languages influenced mathematics and computer science.
  - The Chomsky hierarchy.
  - More classes of languages and grammars have been discovered since 1957.
- Chomsky himself gave up the idea of defining a language as a mind-external set of sentences (e-language) in favor of an approach that seeks to describe the mind-internal grammar of speakers (i-language).
- Many linguists still do believe in the original program, however.
  - Berwick et al 2011 investigate the formal properties of bird song.

The Chomsky Hierarchy

- Finite languages
- Regular languages
  - Grammar rules: X → aY, X → aX, X → a, …
- Context-free languages
  - Grammar rules: X → Ya, X → aXa, X → ab, …
- Mildly context-sensitive languages
- Context sensitive languages
  - Grammar rules: X → Ya/W_Z, …
- Recursively enumerable languages

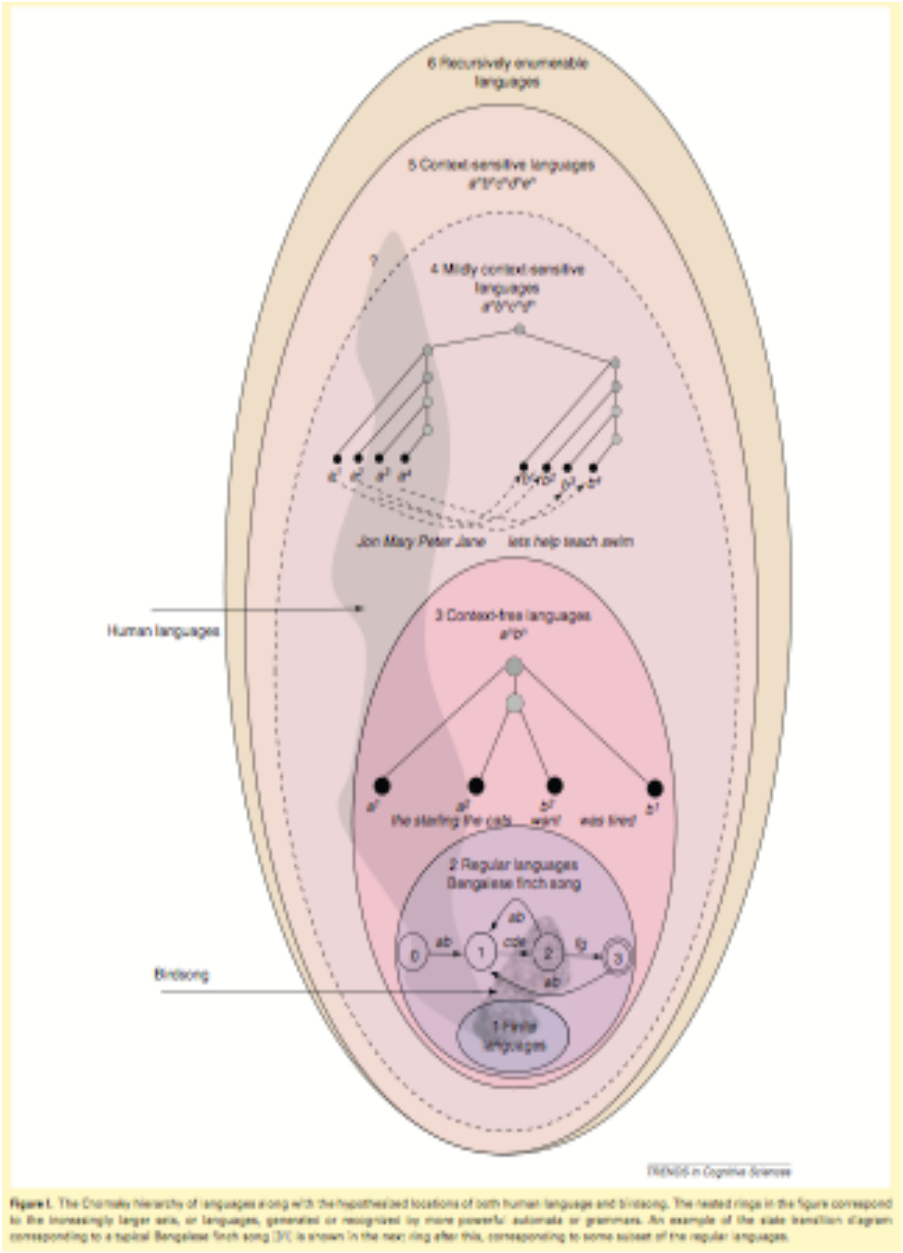


Figure I. The Chomsky hierarchy of languages along with the hypothesized locations of both human language and birdsong. The nested rings in the figure correspond to the increasingly larger sets, or languages, generated or recognized by more powerful automata or grammars. An example of the state transition diagram corresponding to a typical Bengalese finch song [31] is shown in the next ring after this, corresponding to some subset of the regular languages.

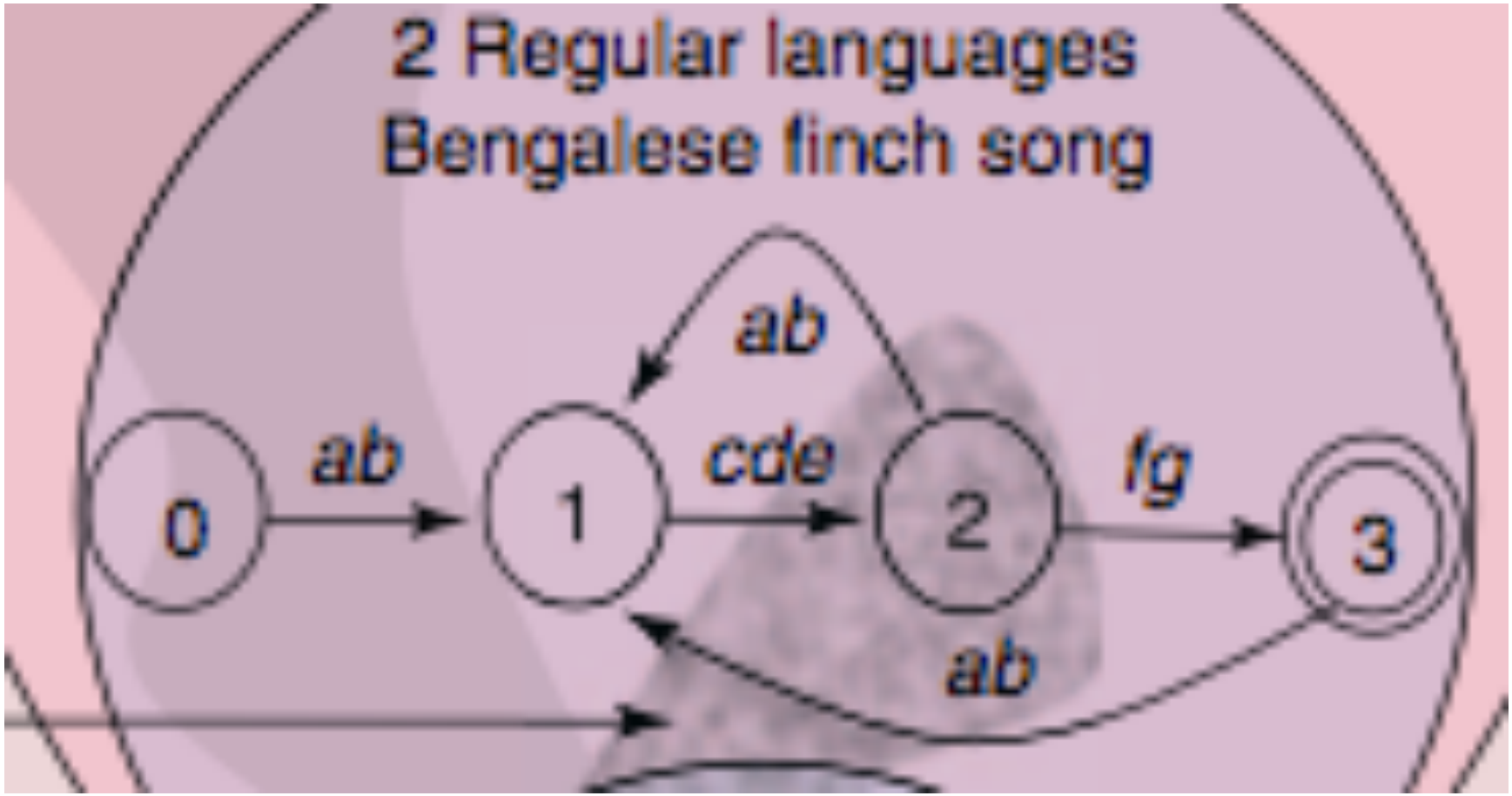
Languages

- Regular:
  - S → aS, S → b
    - *ab, aab, aaab, aaaab, ….*
    - *The old man, the old old man, the old old old man, …*
- Context free: $a^n b^n$
  - S → aB, B → Sb, S → ε
    - *ab, aabb, aaabbb, aaaabbbb, …*
    - *The rat the cat chased ate the cheese.*
    - *The rat the cat the dog bit chased ate the cheese.*
  - Nested dependencies
- Mildly context-sensitive : $a^n b^m c^n d^m$
    - *abcd, aabccd, aaabbcccdd, …*
    - *John, Peter and Bill laughed, smiled, and cried, respectively.*
    - Dutch: *Jon Mary Peter Jane lets help teach swim*
      'Jon lets Mary help Peter teach Jane to swim')
  - Cross-serial dependencies

Automata

- Finite languages
- Regular languages
  - Recognizable by a Finite State Automaton (FSA)
    - A directed graph of states with labeled edges, a finite state transition network.
- Context-free languages
  - Recognizable by a Pushdown Automaton (PDA)
    - Finite-state machines augmented with a potentially unbounded auxiliary memory that can be accessed from the top down.
    - An FSA augmented with the ability to use subroutines, yielding the recursive transition networks.
- Mildly context-sensitive languages
    - A PDA whose stacks might themselves be augmented with embedded stacks.
- Context sensitive languages
- Recursively enumerable languages
  - Recognizable by a Turing Machine

Recording


2 Regular languages
Bengalese finch song
ab
ab
cde
fg
0
1
2
3
ab

Subclasses of regular languages

- Starling songs is locally testable: it can be recognized by bigrams for the most part.
  - The next motif is predictable by the immediately preceding motif.
    - *ab*, *abab*, *abababab*,…
    - Aside from the beginning and end of a song, a bird could check whether the song is well formed by using two bigram templates: [*a*-*b*] and [*b*-*a*].
- However, the Bengalese finch song isn't locally testable.
  - The notes *cde fg* can appear arbitrarily far from either end and bracketed on both sides by some arbitrarily long number of *cde ab* repetitions.
  - Upon encountering the first *cde ab* the FSA must use memory and wait until the required *cde fg* sequence appears.

Birdsong and human phonology

- Bird songs correspond to a subclass of finite state languages (k-reversible) that are easily learnable.
  - Through examination of the local note sequences that can follow from any two states, determining whether the two states should be considered equivalent.
- Human language phonology is also now known to be describable purely in terms of FSA.
  - Phonotactic constraints that determine that *ptak* is not a possible English word, while *plast* might be are also k-reversible.
    - Such constraints are not all-or-none but gradient.
    - Gradient constraints might characterize bird song as well.
    - Adjacent acoustic dependencies of this sort are readily learnable from an early age using statistical and prosodic cues.
  - Vowel harmony rules, as in Turkish, go beyond this simplest kind of strictly local description but remain finite state.

Tweets to phrases: the role of words I

- Birdsong involves complex combinations of sound elements, but the elements are not combined to yield novel 'meanings'.
  - The birdsong conceptual-intentional component is impoverished.
  - It is not connected to a syntactic and externalization component.
  - Song sequence changes alter message strength but not message type.
- Human syntax is intimately wedded to our conceptual system.
  - Involving words in both their syntactic and semantic aspects.
    - 'red' + 'apples' ≠ 'green' + 'apples'
  - It seems plausible that this single distinction drives fundamental differences between birdsong and human syntax.
  - Human language but not birdsong involves compositional semantic creativity.

Tweets to phrases: the role of words II

- Unlike birdsong, human language sentences are potentially unbounded in length and structure.
  - Limited only by extraneous factors, such as short-term memory or lung capacity.
  - Here too words are important:
    - Verb *ate* + Noun *apples* = Verb Phrase *ate apples*, with the properties of its head, a verb.
    - Such word-driven hierarchical structure building can even involve nested dependencies or cross-serial dependencies.
- Birdsong motifs lack word-centric 'heads'.
  - A starling song might consist of a sequence of *warbles* and *rattles*
  - But there is no corresponding way in which the acoustical features of the *warble* class are then used to 'name' distinctively the *warble-rattle* sequence as a whole.
  - So that the combinations could then be manipulated as single-unit phrases into ever-more complex syntactic structures.

Is recursion for the birds?

- Starlings can be trained to recognize languages of the form $a^n b^n$. (context-free)
  - a string of *warbles* followed by an equal number of *rattles*.
  - Getner et al 2006 (*Science)*
- But there is no evidence that the starlings pair up the *warbles* and *rattles* in a nested fashion:
  - $a^1\ a^2\ a^3\ b^3\ b^2\ b^1$.

Birdsong phrase structure?

- Still, birds do seem to have a limited ability to construct phrases.
  - Fodor et al. 1965 showed that a 'click' in the middle of a phrase such as *ate the apples*, tend to be perceived at the beginning or end of the phrase.
  - Suge & Okonoya 2010 showed that Bengalese Finches treat 3-4 note sequences such as *cde* as 'chunks', responding as if the click were at an edge of the chunk.
- Perhaps with the addition of words, humans could label and hold in memory phrases in addition to single words.
  - Once words infiltrated the basic pre-existing syntactic machinery, the combinatory possibilities became open ended.