On the power of rhythm in language and music





Tamara Rathcke

Discovering Linguistics/Linguistic Discoveries
7 November 2016

University of Kent

Rhythm in human communication and interaction

What is rhythm?

Let's look at Wikipedia:

Rhythm (from Greek ῥυθμός, rhythmos, *"any regular recurring motion, symmetry",* Liddell and Scott 1996) generally means a *"movement marked by the regulated succession of strong and weak elements, or of opposite or different conditions"*. (given examples: dance, music and language)

- strong and weak elements (~prominence alternations)
- regularity of their occurrence (in time)

Finger-tap or footfall to rhythm

- In electronic music without vocals (DJ Shadow)
- In instrumental music with vocals (The Beatles)
- In English nursery rhymes (Glasgow accent)
- In spontaneous/unscripted speech (Newcastle accent)




Regularity and isochrony in language

Languages do have rhythm and the required regularity (even isochrony) but this regularity occurs at different levels, depending on the language (Lloyd James 1940, Pike 1945, Abercrombie 1967)

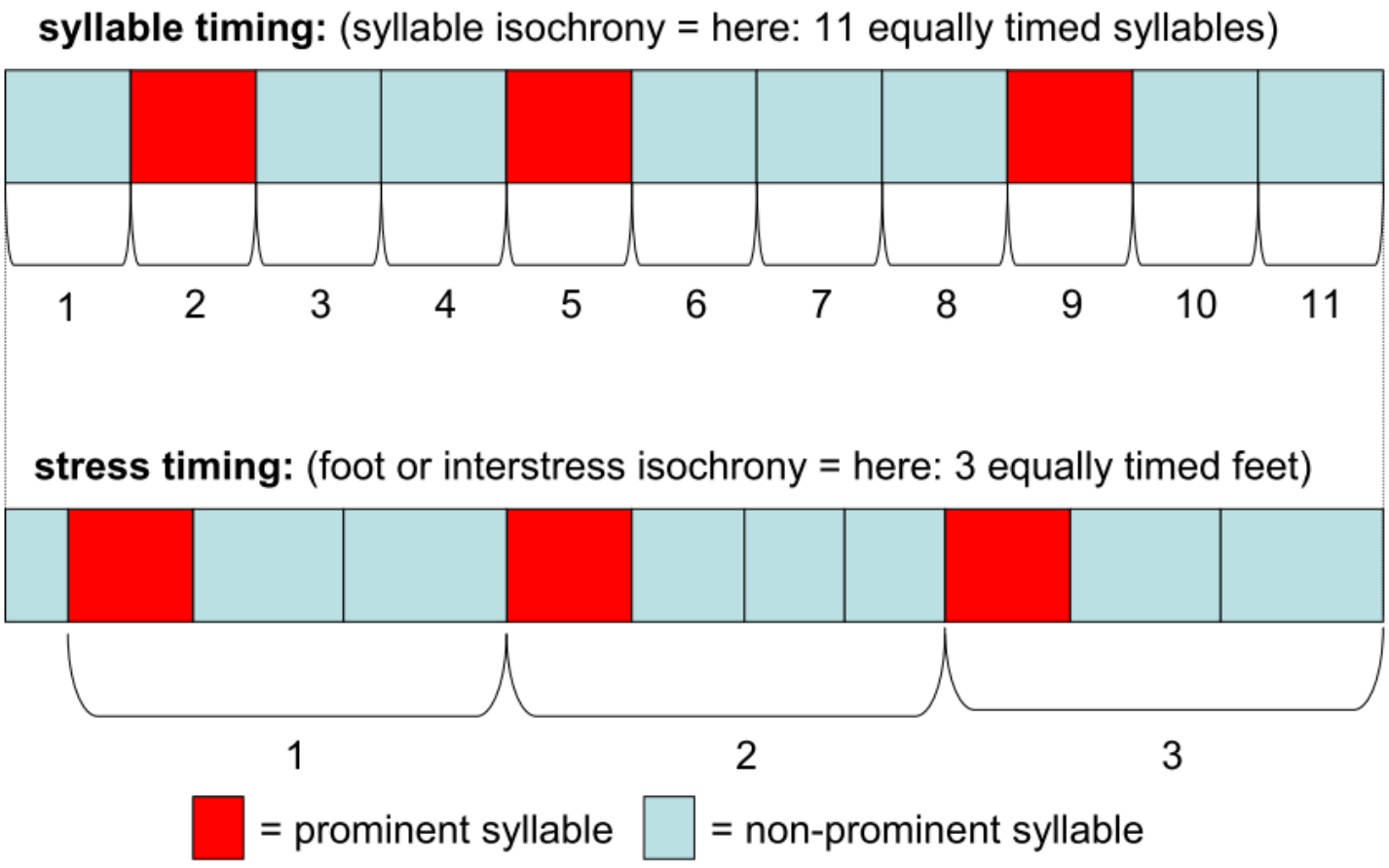
"syllable-timed"

- *"machine-gun"* rhythm
- Examples: Spanish, French
- isochrony of syllables, i.e. each syllable has equal duration

"stress-timed"

- *"morse-code"* rhythm
- Examples: English, Dutch, German
- isochrony of stressed syllables, i.e. each inter-stress interval has equal duration

The early theory of ‘rhythm class’



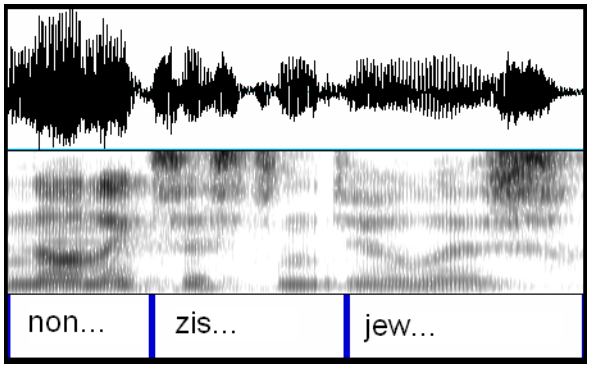
Empirical validation of ‘rhythm class’

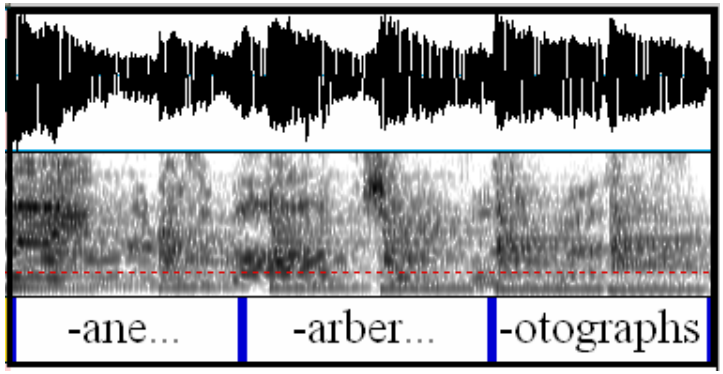
There is ***no isochrony*** in speech signals (cf. tapping to music vs. rhyme/spontaneous speech):

- Syllable, mora and inter-stress durations are extremely variable

the |**non**-ek| **zis**tent |**je**weller’s

In Penny |**Lane** there is a |**bar**ber showing |**pho**tographs





≈ 0.3, 0.4, 0.5 s

≈ 1.1s

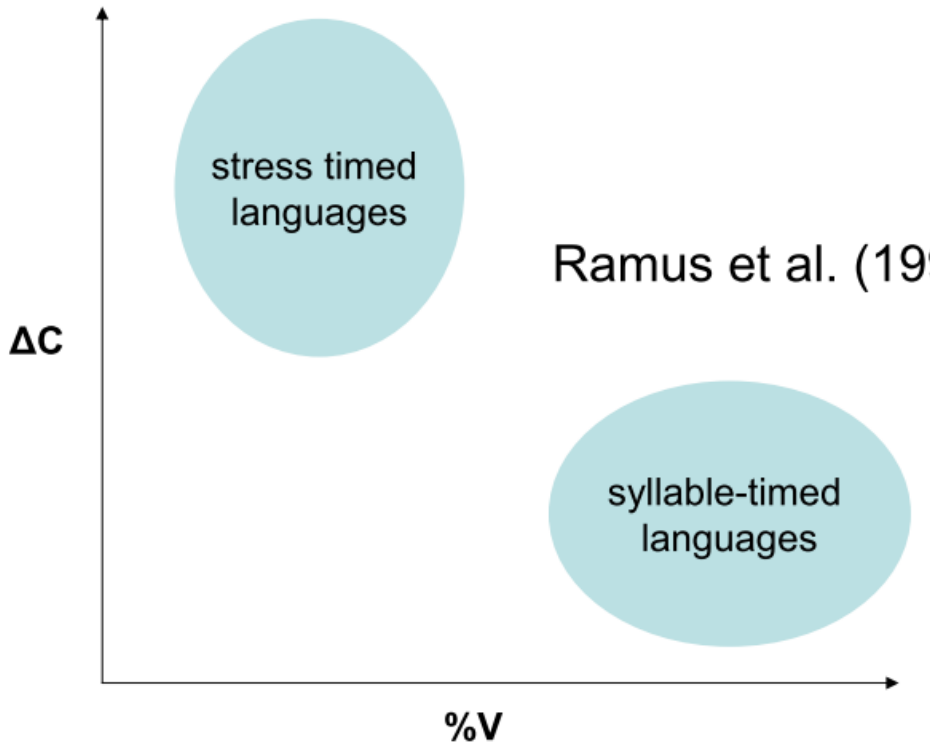
The late theory of ‘rhythm class’

- Revival of interest in rhythm classes since 1990s
- There is a correlation between rhythm class and use of consonants/vowels (Dauer 1983/1987):
  - *Syllable-timed:* no complex consonant clusters, no vowel reduction in weak syllables → longer vocalic portions, less durational variability in C and V
  - *Stress-timed:* complex consonant clusters, vowel reduction (e.g. *strengths* CCCVCCC) → shorter vocalic portions, greater durational variability in C and V

→ **‘rhythm metrics’**

Rhythm metrics

Languages from the two rhythm classes can be well discriminated using a series of “rhythm metrics”



Ramus et al. (1999) findings

Problems with the metrics and the idea of rhythm class:

Arvaniti (2010, 2012), Rathcke & Smith (2015a,b)

(%V = percentage of vocalic intervals

ΔC = standard deviation of consonantal intervals)

Is the rhythm dichotomy real?

- Original evidence in favour of ‘rhythm class’ distinction comes from discrimination experiments: Ramus et al. (2003) tested language discrimination using *‘sasasa’* stimuli
  - only languages from different classes are discriminated (Japanese and English) but languages from the same ‘rhythm class’ are not (English and Dutch)
- Counter-evidence:
  - Language discrimination results are just an artefact of speech rate differences in the stimuli (Arvaniti & Rodriguez 2013)

Further counter-evidence

…comes from identification experiments:

- naïve participants cannot use the two categories (Arvaniti & Rodriguez 2013; Miller 1984)
- phonetically trained listeners do not produce consistent patterns (Miller 1984)
- expectation (~knowledge of the dichotomy) of trained listeners is important (Rathcke & Simith 2015)

‘Rhythm class’ is not the (whole) story of language rhythm

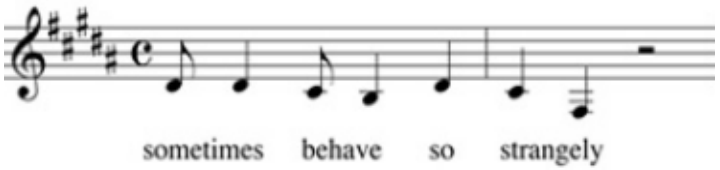
- Rhythm metrics look only at the relative timing of single sounds neighbouring within a phrase!

However,

- rhythm is a feature of *prosody*
  - long-distance, more than just a segment
- rhythm connects language and music
  - from music perspective, language rhythm can be defined as a systematic patterning of units (syllables, words) in terms of *prominence* (~strong/weak), *grouping* and *timing* (cf. Patel 2008)
  - ‘the speech-to-song illusion’ (Deutsch 1993)

What is the Speech-To-Song Illusion?

D. Deutsch 1995



Repeated several times in exactly the same form, a *spoken* phrase shifts to being heard as *sung*

- Tierney, Dick, Deutsch & Sereno 2013: audio books

Why do we experience the illusion?

- *"This illusion is in line with what philosophers and musicians have been arguing for centuries, that strong linkages must exist between speech and music…"* (Deutsch 1995)
- No different sets of physical properties that are unique to speech or song (Deutsch et al. 1995, 2008, 2011)
  - Is this really the case? Is the illusion perception unaffected by the acoustic properties of the phrase? (Falk, Rathcke & Dalla Bella 2014)

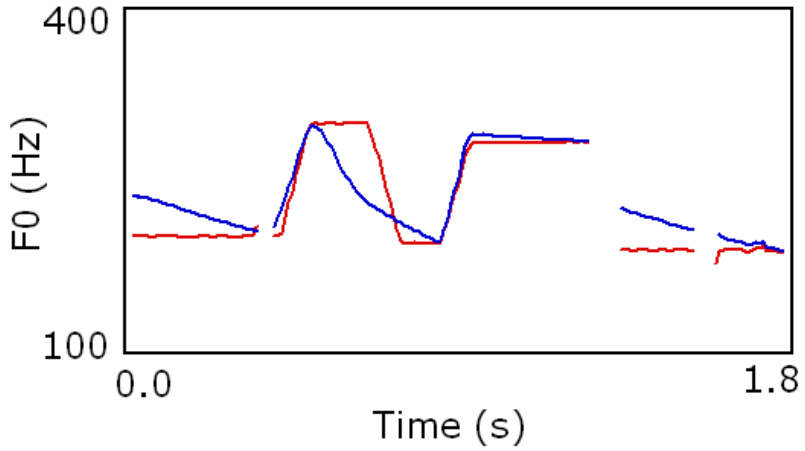
    Hypothesis: The more song-related acoustic properties (tonal, rhythmic) are integrated in a signal, the easier/ quicker the shift

    - Tested with 92 listeners of German

Stimuli: Tonal manipulations

1. Global pitch shape

variability vs. stability in the pitch contour



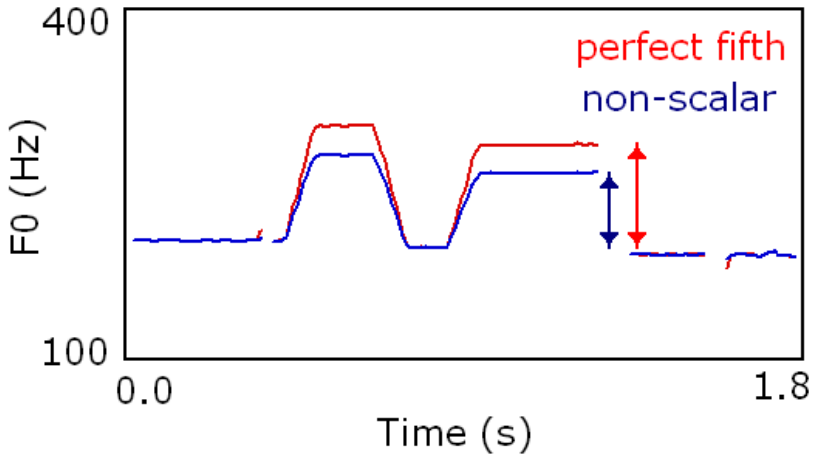
Speech-like

Song-like

Stimuli: Tonal manipulations

2. Local pitch intervals

non-scalar vs. scalar relationships between pitch targets



Speech-like

Song-like

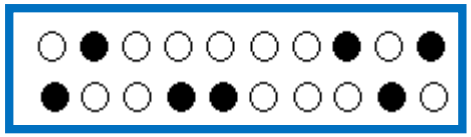
Stimuli: Rhythmic features

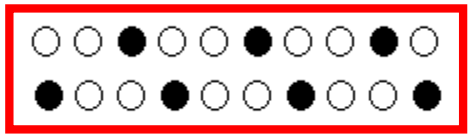
1. Linguistic rhythm

   s/w-alternations within a sentence containing 10 syllables

Exp. 1

Exp. 2





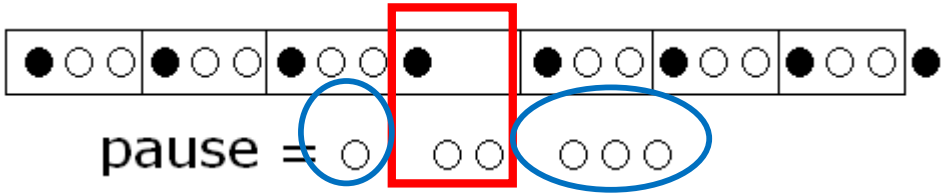
Speech-like (irregular)

Song-like (regular)

Stimuli: Rhythmic features

3. Macro-rhythm

*regularity of **s** (=strong beats) throughout the loop*

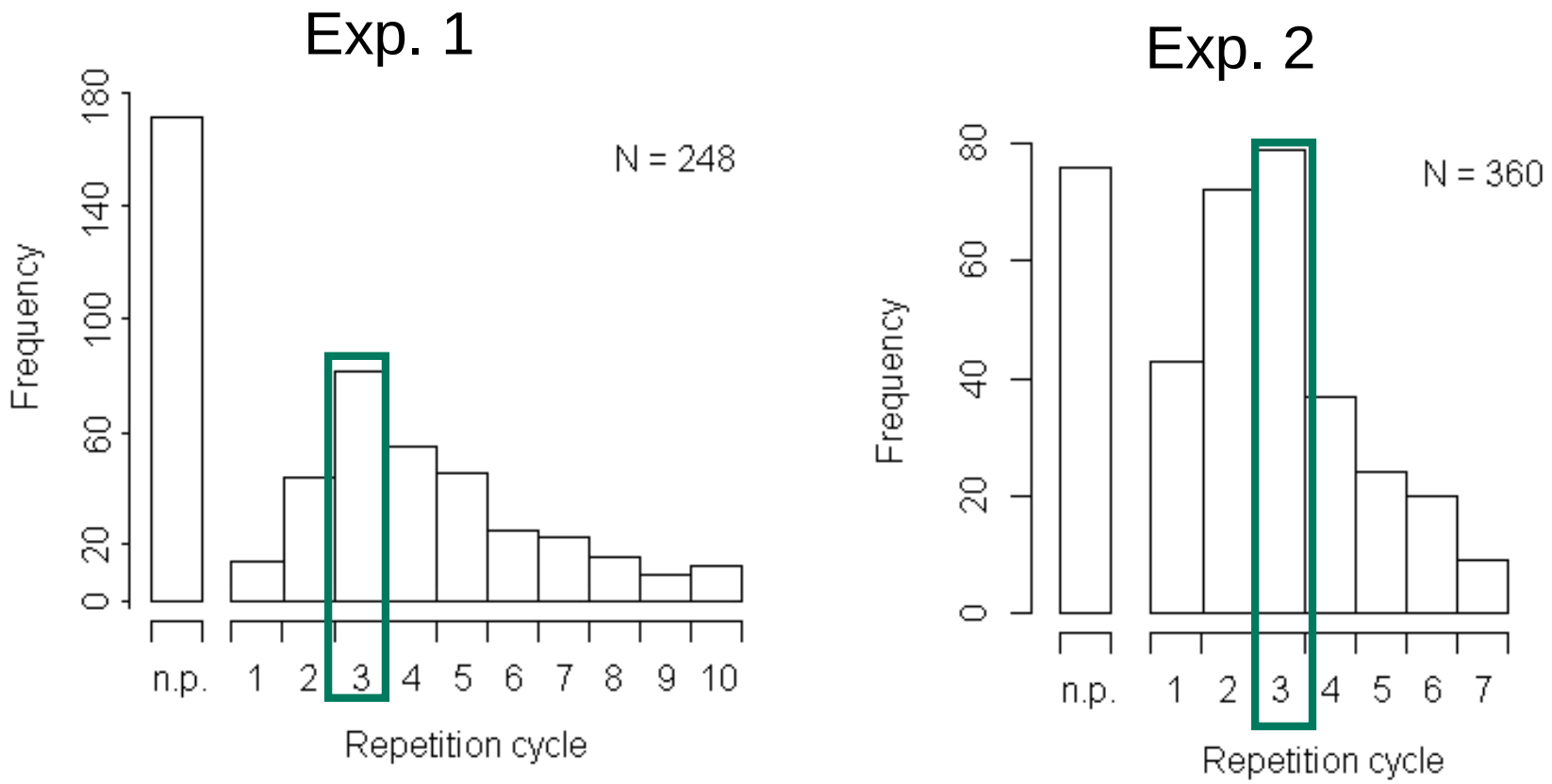


shortened / matched / lengthened

Speech-like

Song-like

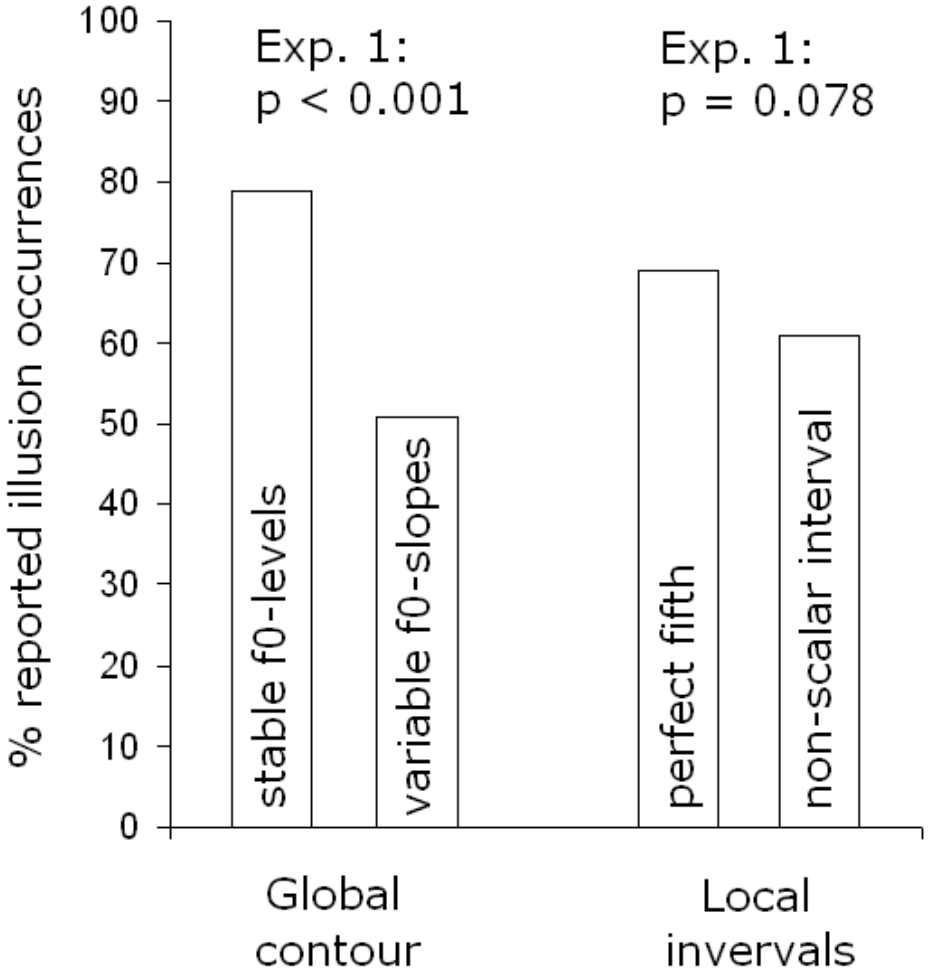
Overall pattern of results

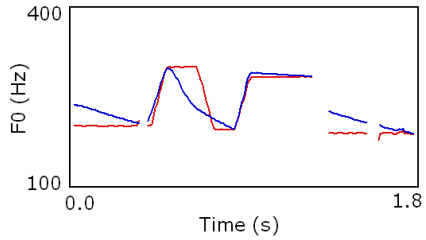


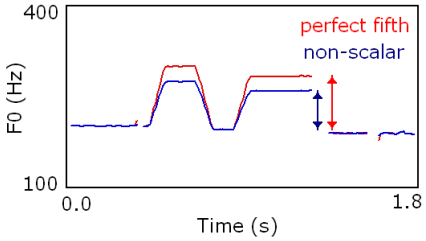
n.p.= no illusion perceived

Repetition cycle = illusion perceived during N repetition within the loop

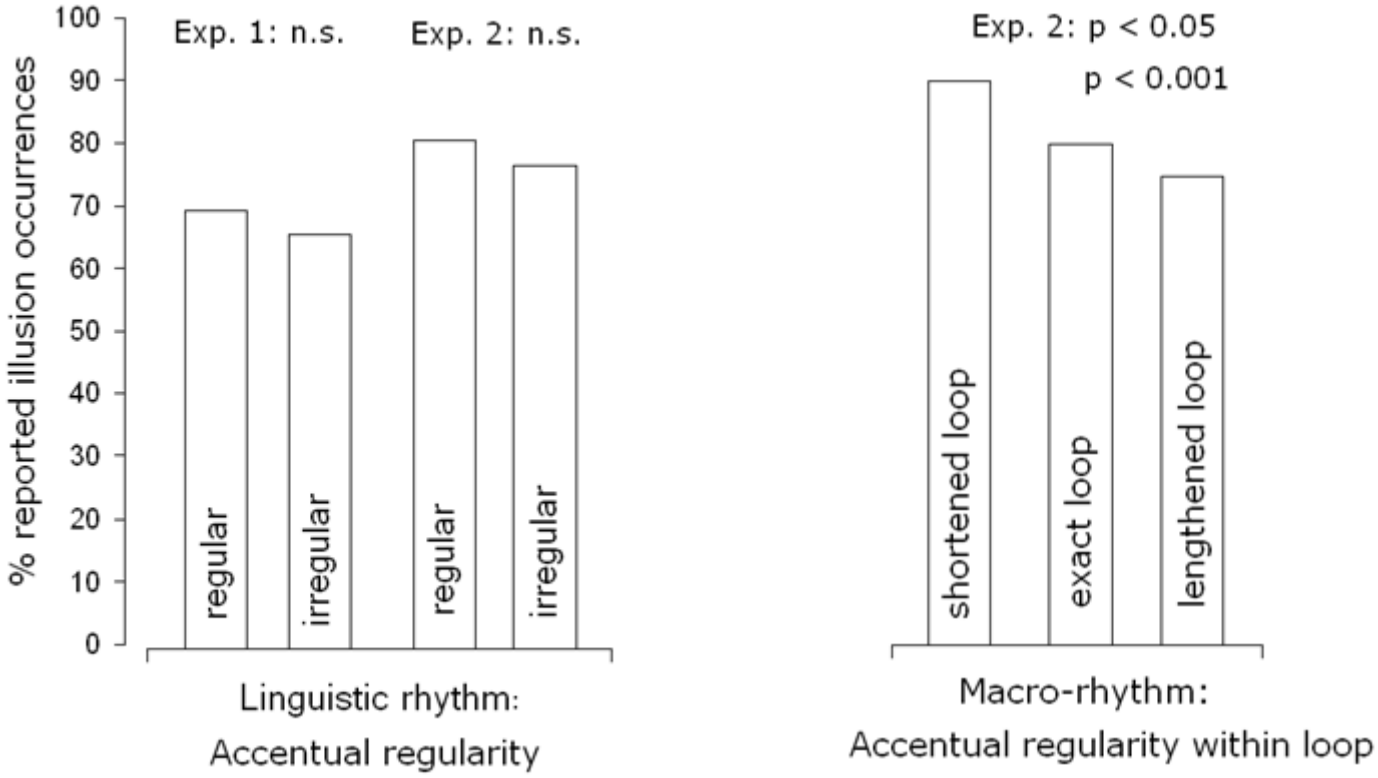
Tonal cues


Exp. 1:
p < 0.001
Exp. 1:
p = 0.078
% reported illusion occurrences
100
90
80
70
60
50
40
30
20
10
0
stable f0-levels
variable f0-slopes
perfect fifth
non-scalar interval
Global
contour
Local
invervals


400
F0 (Hz)
100
0.0
1.8
Time (s)


400
perfect fifth
non-scalar
F0 (Hz)
100
0.0
1.8
Time (s)

Rhythmic structure



- Isochrony does not matter for the illusion! The change to the perception of song can occur regardless of the spoken phrase being more or less regular.

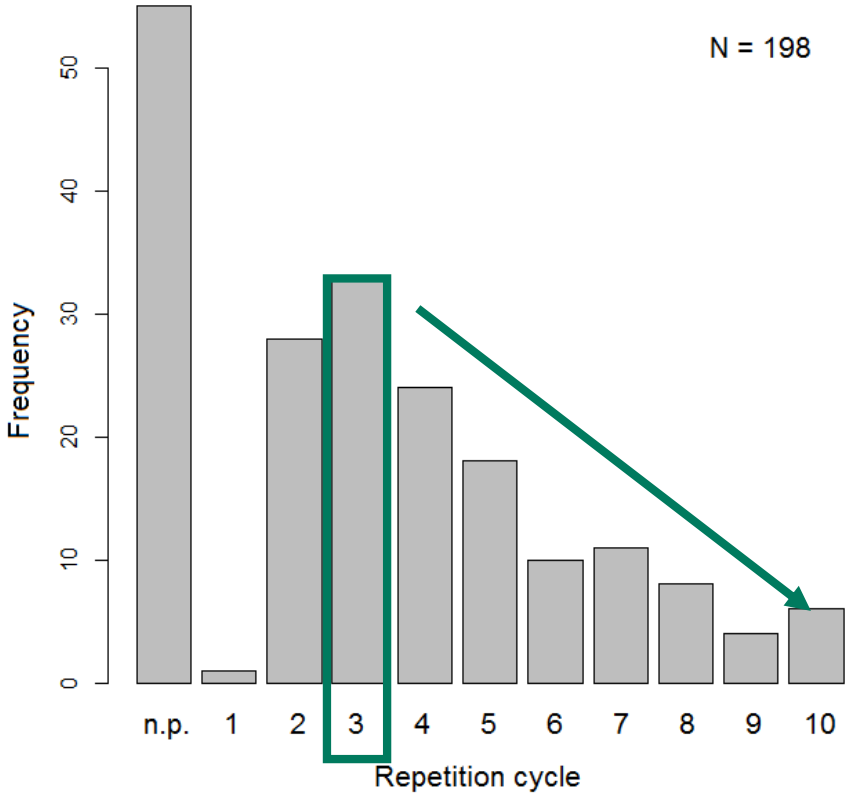
Why does the illusion happen?

- Perceptual effects of massed repetitions
  - verbal transformations (Warren 1961)
  - semantic satiation (Smith 1984)
  - syntactic fatigue (Francom 2009)
- The speech-to-song shift:
  - **meaning decay**
    (cf. 3 repetitions in semantic satiation)
  - **rhythm prevalence**
    (through recurrent patterning of strong/weak beats)
  - **melodic re-analysis**
    (given appropriate acoustics ~ *"a group of tones in love with each other"* A. Patel 2007)

Perception of the illusion in non-native listeners

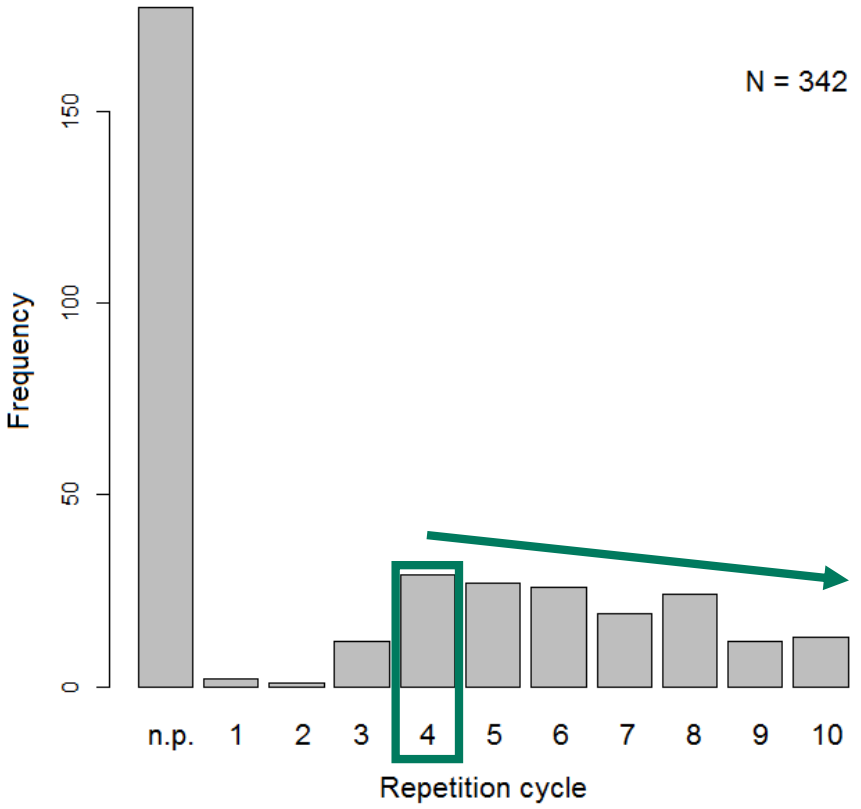
If there are delays in the immediate access to the meaning of the phrase, the shift will be delayed or will not happen at all.

28%



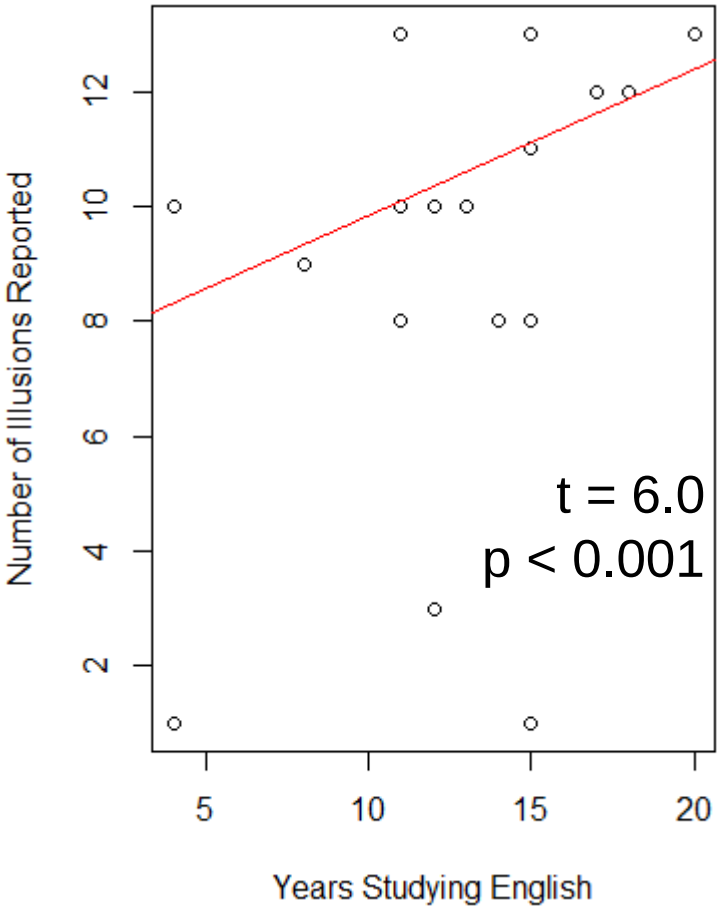
Native listeners

52%



Non-native listeners

Role of L2 proficiency



No effect of

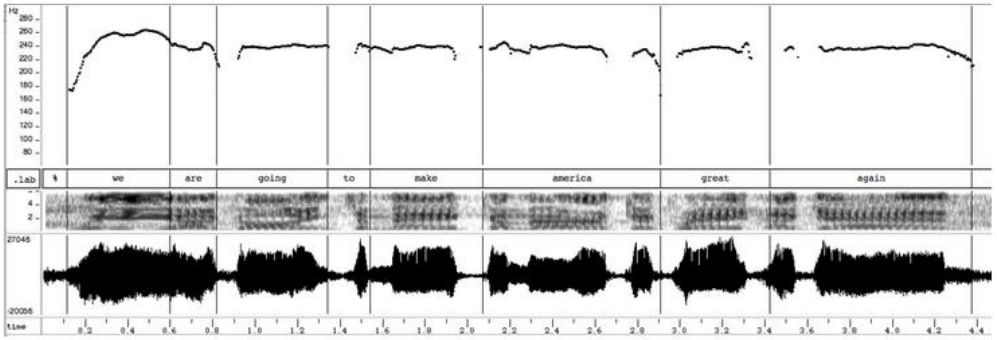
- self-reported levels of proficiency in English
  - scale between 1(poor) to 10 (native-like)
- number of years resident in the UK
  - varied between 3 and 19

Not many languages studied (so far, English, German and Chinese), currently work on English and French (Rathcke, Falk & Dalla Bella)

Crossing the boundaries between music and language

- Spoken and repetitive signals in music:
  - DJ Shadow

- Singing/chanting in speech:
  - Donald Trump (the opening phrase of his speech in Mannheim PA, 1/10/2016, cf. Mark Liberman) – a flat contour at 240 Hz



Rhythm in political speech

Speech accompanying gestures

4 types of gestures

1. *denotative* (**DG**): refers to a feature of the action described ("*he climbed up the stairs*")
2. *connotative* (**CG**): form of gesture is metaphoric/representational ("*the meeting went on and on…"*
3. *indexical* (**IG**): helps locate something in the physical space ("*that document shows*")
4. *rhythmic* (**RG**): baton-like, beat gestures, emphasise the discourse, no information

Speakers

Donald Trump

- DOB: 14/06/1946
- businessman, TV personality, author, politician, nominee of the Republican Party for the US President election in 2016
- L1: American English (NY City)

~25 min material per speaker (2 shorter speeches for DT, one longer for AS); downloaded from YouTube

Arnold Schwarzenegger

- DOB: 30/07/1947
- businessman, actor, producer, bodybuilder, politician (Republican), Governor of California 2003-2011
- L1: Austrian German; L2: American English (moved to the US in 1968, aged 21)

This Analysis of Donald Trump's Speech Patterns May Explain His Unlikely Rise: WATCH

by **Sean Mandell**
January 5, 2016 | 1:56pm

Trump shows speech can have effect without making complete sense

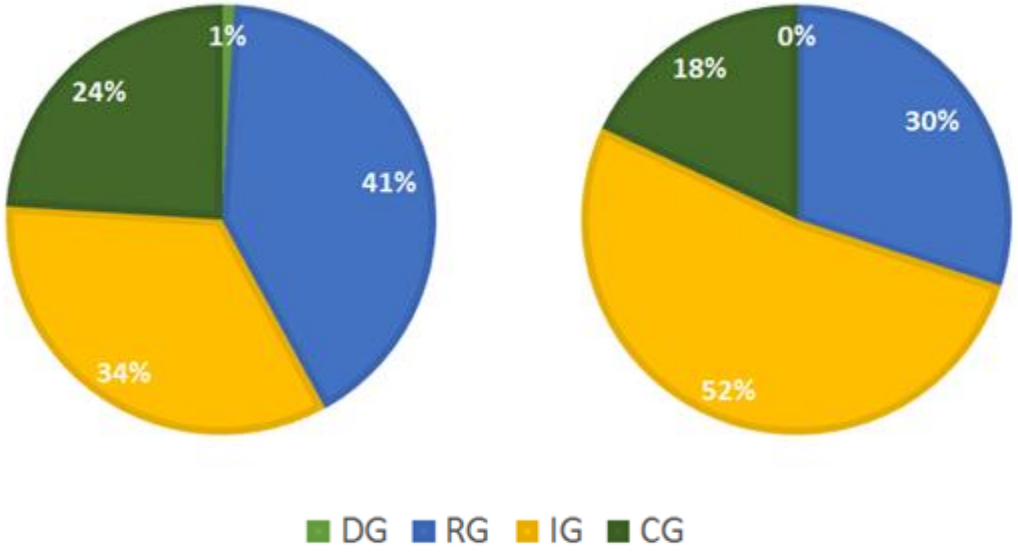
The billionaire used conventional tools of the orator to work up his presidential speech

The Art of Persuasion

Is Donald Trump's 'broken speech' the key to his success? Linguists say strange patterns may make him more 'authentic, relatable and trustworthy'

- Trump's speech is appealing because it contrasts with rehearsed style
- Repeating phrases can be effective to show you're a consistent candidate
- Research finds most candidates show grammar at 6-8 grade level
- Lincoln showed grammar skills at the 11th grade level
- For more of the latest on Donald Trump visit www.dailymail.co.uk/trump

Results – the use of rhythmic gesture



- *denotative* (**DG**): "*he climbed up the stairs*"
  - *rhythmic* (**RG**)
  - *indexical* (**IG**): "*that document shows*"
- *connotative* (**CG**): "*the meeting went on and on…*"

Results – the use of rhythmic auditory features

Feature	DT	AS
Mean speech rate	4.0 syll/sec	4.0 syll/sec
Mean pitch range	18 st	13 st
Emphatic stress use	3.92 times/min	2.59 times/min

- Overall, Donald Trump is a skillfull user of audio-visual rhythmic features in his speech (rhythmic beat gesture, clear rhythmic structure through the use of emphatic prominence)
  - Not studied here: repetition of phrases (local rhythmic feature), repetition of syntactic structures (long distance rhythmic feature) – preliminary observations
- Repetition, and also an extensive use of gesture, enables an easier cognitive processing as less cognitive resources are engaged (REF)

Functions of speech rhythm

- Establishment of a rhythmic pattern guides *attention*; allows prediction of when key information will occur, e.g. lists:

  po**ta**toes, to**ma**toes, **o**ranges, **ca**rrots and **bread**

  vs.

  oranges, po**ta**toes, **bread**, **ca**rrots and to**ma**toes

- Accentuation & phrasing help to make speech *easy to understand*, and help with conversational turn-taking
  - If the overall rhythmic pattern is wrong, speech is hard to understand (e.g. synthetic speech; foreign-accented speech)

Functions of rhythm in general

- Rhythm recognition creates *a state of pleasure*
  - Pre-linguistic children enjoy rhythm (M. Zentner & colleagues)
- Allows *synchronised action* through a unified temporal structure:
  - chanting, singing, dancing, bouncing a baby
- Enhances *social bonding*
  - Headphone parties: we feel stronger connection to and more sympathy for the people dancing to the same music (I. Cross & colleagues)
- Implicates *a survival advantage* within a community:
  - Individuals who walk or sing together are subsequently more helpful, compliant or cooperative with each other (Wintermuth and colleagues)
  - As early as in 4-year-olds (Tomasello and colleagues) or even younger (Cirelli and colleagues)

Open questions in linguistic rhythm research

- There is no definition of rhythm that researchers would agree on (see Cummins 2012 for a discussion).
- Well evidenced that the rhythmic typology stress-timing/syllable-timing is not valid, though it is unclear if a rhythmic typology is possible at all.
- What are the best methods to study rhythm?
- Mostly European languages (lexical prominence: strong/weak alternations) studied to date. Korean, French – no concept of lexical stress. Does the concept of rhythm still apply to these languages?
- Research to date has concentrated to auditory features of rhythm, the role of visual cues is still poorly understood

Thank you!

Questions?
tvr@kent.ac.uk





University of Kent