University of Essex

Mental Representation of Tone in Bemba

Nancy C. Kula
University of Essex
nckula@essex.ac.uk

TAL 2018, Beuth University, Berlin, 18-20 June

Outline

- Introduction
- Bemba Tone
- Research Questions
- Experiment 1: Production
- Experiment 2: Perception
- Experiment 3: Follow-up
- Findings
- Conclusions

Previous studies

- Psycholinguistic studies on tone processing in Asian languages argue that tone is processed in a similar way to vowels and consonants and is thus an integral part of the representation of a segment (Braun & Johnson 2011, Repp & Lin 1999, Ye & Connine 1999, Sun & Huang 2012)
- This cannot be assumed for Bantu where we have a few lexically specified tones that undergo tone spreading processes (Derived tone)
- Bantu: 2-3 tones H, M, L, relative tone, can have contours

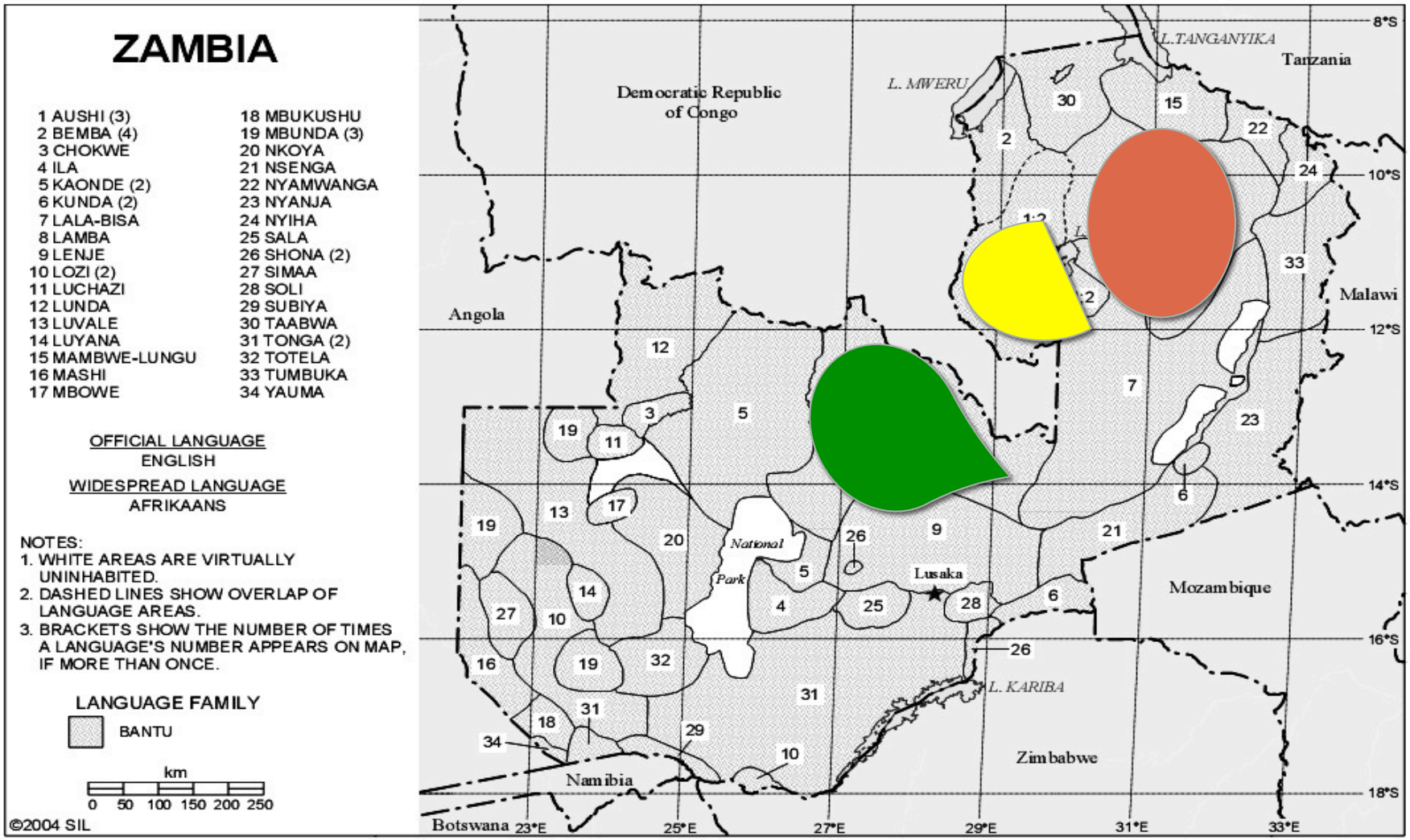
Focus: Bemba tone

- Discussion here is based on Bickmore & Kula (2013), Kula & Bickmore (2015)

- Other seminal works on Bemba tone: Guthrie (1945), Sharman & Meeussen (1955), Sharman (1956), Mann (1977), Philippson (1998)

- Bemba is a central Bantu language (M42), spoken mainly in Zambia but also in parts of the DRC, in Zambia about 2.3 million speakers and similar number of second language speakers

3 main dialects



- Northern Bemba

- original inhabitants
- main dialect
- spoken amongst other Bemba dialects

- Copperbelt Bemba

- descended from NB
- close relations with Town Bemba
- spoken among other Bantu languages

Bemba basics

Vowel length is contrastive:

a. ú-kú-shík-á ‘to be deep’
b. ú-kú-shíík-á ‘to bury’

TBU is the mora. Attested syllables: Cà, Càà, Cá, Cáá, Cáà

Verbs exhibit a tonal contrast on root-initial TBU

a. lùk-á ‘weave!’
b. lúk-á ‘vomit!’

Tones can be grammatical, introduced by TAMs (Melodic Highs)

NB & CB: unbounded H spreading

- A high tone undergoes unbounded spreading if no other H follows, or the word is phrase final

bá-ká-fík-á	‘they will arrive’
bá-ká-bíl-á	‘they will sew’
bá-lá-bóómb-á	‘they work’
bá-lá-lóóndólólá	‘they explain’

Bounded spreading

- A high tone undergoes bounded spreading if there is another H following, or if there is a following constituent
- NB: Binary

bá-ká-fik-a kó	‘they will arrive there’
bá-ká-bil-a kó	‘they will sew a bit’
bá-lúb-ul-ul-é	‘they should explain’

- CB: Ternary

bá-ká-fík-a kó	‘they will arrive there’
bá-ká-bíl-a kó	‘they will sew a bit’
bá-lúb-úl-ul-é	‘they should explain’

Ternary spreading

- Problematic for phonology
- Violates locality: the idea that a trigger and a target must be adjacent
- Goes against assumptions of binary branching structure

- In Bickmore & Kula (2013) we argue that ternary spreading is achieved by two distinct processes:

- **High Doubling**: spreads a H a single mora to the right

- **Secondary High Doubling**: further spreads a H to the first mora of the following syllable

- Other cases of ‘ternary’ spread in Bantu:

(a) Shona dialects (Myers 1987, Hewitt & Prince 1989, Odden 1981, Johnson & Paster 2012)
(b) Ikalanga (Hyman & Mathangwane (1998)
(c) Tswana (Creissels 1998)
(d) (Dagbani (Gur language, Hyman 1993))
(e) Ternary displacement in Sukuma (Goldsmith 1985, Sietsema 1989, Batibo 1991, Roberts 1992 inter alia)

Is ternary spread phonological or phonetic?

- Is it a case of tonal allophony, just a way of realising bounded spread in CB, i.e. it is not specifically ternary
- Is it a case of phonetic overshoot?
- If ternary spread applies to nonce words this suggests inferential, rule-based processes (Pinnow & Connine 2014)

- **Research Questions**:
- What is the mental representation of ternary spread?
- Do speakers specifically process ternary spread?
- What is the representation of derived tone?
- What is the representation of a relative two-way tone contrast?

Two main experiments

- Experiment 1: Production

  Focused on ternary spread

- Experiment 2: Perception

  AX discrimination task (ternary v. binary spread)

- Experiment 3: Perception follow up

  Chinese and German participants

Experiment 1: Production

- Participants
- 23 native speakers of Bemba
- From two tertiary institutions: Nortec College (Ndola) and Nkrumah University (Kabwe)
- Aged between 18 and 66 years (*m*=37.7 years, SD=14.3 years)
- 15 male, 8 female
- No prior training in tone/intonation and were unaware of the purpose of the experiment
- Participated for a small fee

- Participants originated from the 3 main Bemba speaking areas (Luapula, Northern, Copperbelt)
- Had all spent a significant amount of time in the Copperbelt area (ternary dialect: between 1-34 years, *m*=16.5 years, SD=10.9 years)
- None resided in the northern area (binary dialect) at the time of testing

- Two conditions: real and nonce word conditions
- 8 stimuli for each condition
- Each item consisted of a trisyllabic verb stem (e.g., sakula, ‘comb’)
- 4 additional stimuli were used for the familiarization phase (2 nonce, 2 real)
- 10 filler items (8 real, 2 nonce) of 4-syllabic words
- All existing verb stems were lexically low-toned
- Forms were in the habitual aspect with lexically low-toned 1pl subject marker (tu-) followed by *lyonse* ‘all the time’

Stimuli shape

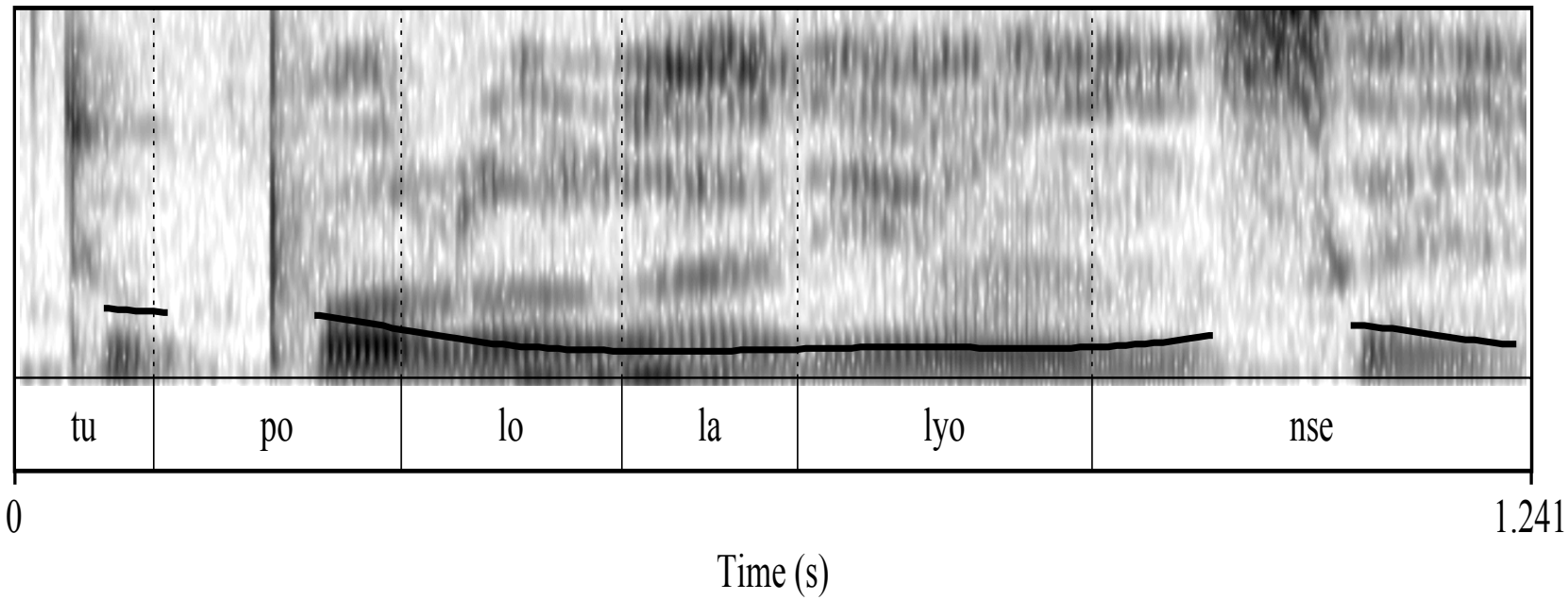
- tu-sakula lyonse ‘we comb all the time’
- tu-safuka lyonse ‘we *verb* all the time’

- Items were recorded with Zoom H4 (44.1kHz, 16Bit)
- The procedure was based on a modified Wug test (Berko, 1958; Ratner & Menn, 2000)
- No pictures were used

Acoustic analyses of the real and nonce words showed no difference in pitch on the tri-syllabic verb stem

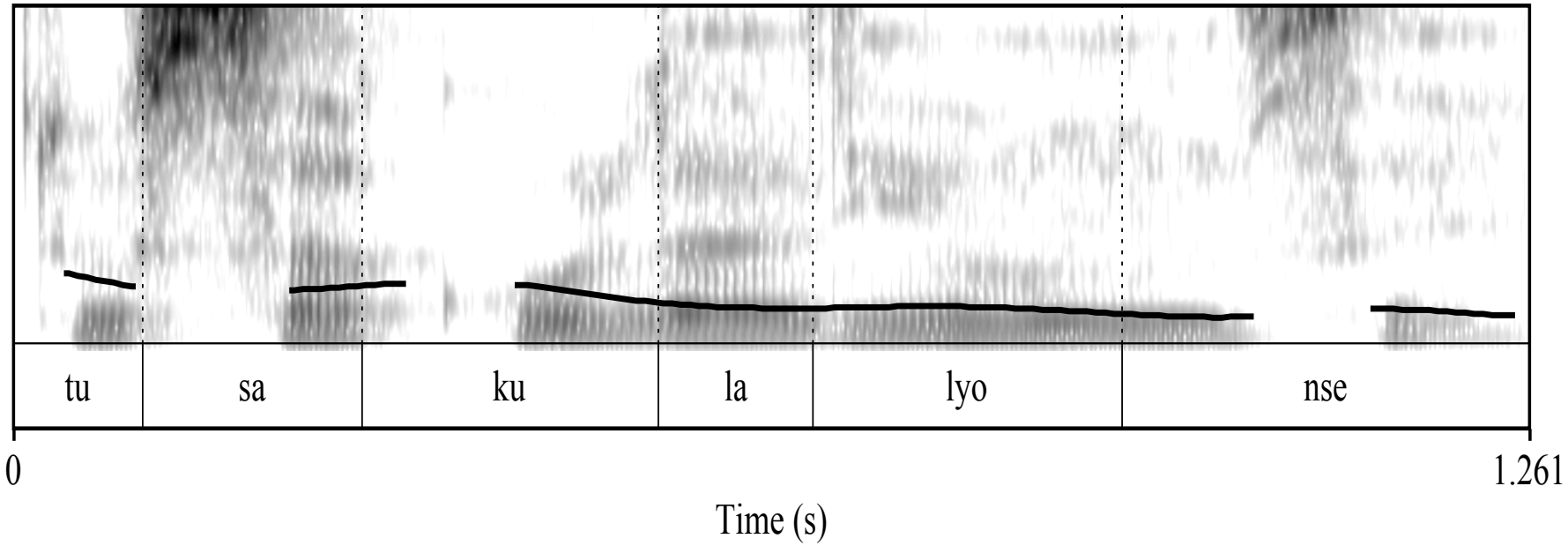
	Average f0 (and SD) in			
	first syllable	second syllable	third syllable	fourth syllable
real stem	142.1	132.4	126.0	119.7
	(4.5)	(4.6)	(5.3)	(3.8)
nonce stem	133.4	127.0	124.1	117.0
	(11.2)	(8.1)	(8.9)	(7.8)
Difference	$p < 0.05$	n.s.	n.s.	n.s.

Nonce word



F0-range is shown between 100 and 300 Hz

Real word



F0-range is shown between 100 and 300 Hz

Task

- Replace the low-toned subject marker *tu-* with the high-toned subject marker *bá-* (class2 SM) and retain the rest of the verb form
- Add the adverb *lyonse after the verb*
- Target words: *básakula lyonse (they comb all the time)*

- This context with a constituent following the verb was intended to induce bounded spreading, i.e. predominantly ternary spread.

Method

- Participants were tested individually in a quiet room
- They were seated in front of a laptop computer and were recorded using a Zoom H4 recorder (44.1kHz, 16Bit)
- They were instructed orally in English and in Bemba
- Each of the verbal forms for the wug task was set up using power-point. They saw an image of a loudspeaker on the screen
- Sound files were played via headphones
- There was no predefined time interval in which participants had to respond but they generally responded within 1000ms
- Participants were allowed to listen to the recording a second time if they wished (~10%)
- The experiment was preceded by 4 practice trials
- The experiment lasted about 12 minutes

Results

- Tonal realisation on the target 3 syllables were annotated by two phonetically trained annotators (one unaware of the purpose of the experiment and other a native speaker)
- Items were classified as unbounded, ternary, binary
- Unbounded spreading was noted to only have occurred when there was a pause between the verb form and the following adverb

	Prosodic realization			
Stem type	Unbounded	Ternary	Binary	Other
Real	5.7	88.9	4.2	1.2
Non-word	5.7	90.0	4.1	0.2

Average percentages of different types of prosodic realizations for real and nonce word stems

Production task findings

- Participants were able to extend their spreading pattern to nonce words
- From an input that consists of all low tones participants were able to produce novel forms with a high-toned initial
- Since there was no high tone at all in the input the high tones after the high-toned subject marker can only be interpreted as the result of tonal spreading
- Therefore ternary spread is phonologically real at least in production

Experiment 2: Perception

- Tested whether native speakers of Bemba have different phonological representations for nonce-words with a binary or ternary spread (Kula & Braun 2013, 2014)
- Conducted a speeded AX (same-different) task with a 2000ms inter-stimulus interval argued to tap into phonological representations (Babel & Johnson 2010, Cowan & Morse 1986, Crowder & Morton 1969)
- Participants were identical to those tested in Experiment 1. All participants first completed Experiment 1 and then Experiment 2.

Materials

- Contrasted binary v. ternary spread in nonce words
- As a control condition we also had the contrast with real words:
- - participants should perform well if they relied on dialect discrimination
- - participants should show a lexical effect if they do not rely on dialect discrimination
- A further control condition contrasted a similar tonal difference involving mid tones but which is not phonologically contrastive

4 conditions

- 1a: binary vs. ternary spread (2H+3L vs. 3H+2L) with a real first syllable with H tone + nonce stem
- 1b: binary vs. ternary spread (2H+3L vs. 3H+2L) with real words

- 2a: 2High-2Mid vs. all-High (real prefixes + nonce stem)
- 2b: 2High-2Mid vs. 2High-2Low (real prefixes + nonce stem)

- Bin v. Ter real prefix + nonce stem *phonologically*
- Bin v. Ter all real word *contrastive*

- HM v. HH real prefix + nonce stem *not phonologically*
- HM v. HL real prefix + nonce stem *contrastive*

stimuli

- 8 stimuli each for real and nonce words (B-T condition)
- 16 nonce words as fillers
- The two real prefixes were a SM and a TAM marker
- SMs: a- (3SG.SM, used 4 times) or ba- (2SM, used 4 times)
- TAMs: -ka- (FUT3, used 4 times) or -la- (HAB, used 4 times)
  Nonce stem had three syllables (e.g. -tekeba) followed by a real post-verbal clitic -ko

- a-ka-tekeba ko ba-ka-tekeba ko
- a-la-tekeba ko ba-la-tekeba ko

- Stimuli for HM-HH and HM-HL had the same two real prefixes (a- and -ka-, four times in each of the HM-HH and the HM-HL conditions)
- Stems in these 16 stimuli were always nonce-word stems (e.g. fusa)

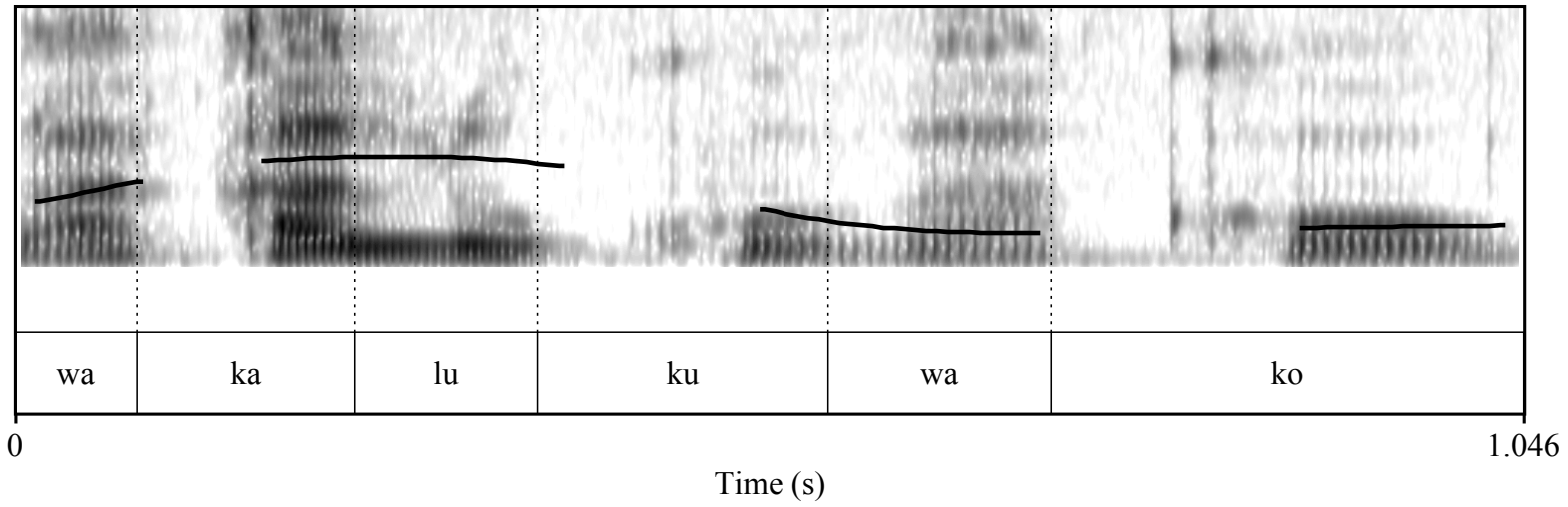
- Example of HM-HH stimuli:
- a-ka-fus-a vs. a-ka-fus-a

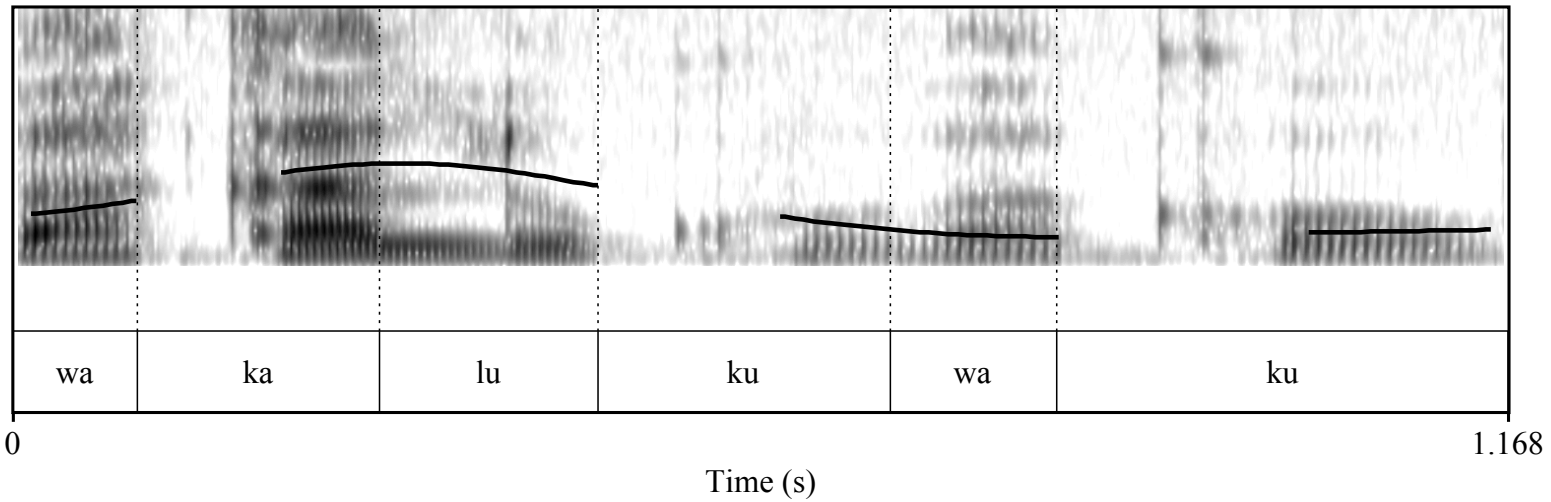
  H H M M H H H H

- 16 fillers had nonce prefixes and stems (e.g., falimasa): all high, all low or a declining high-to-low pattern
- There were 4 practice trails

- All experimental items were produced twice in isolation by a female native speaker of Bemba
- A single speaker was used for all stimuli to avoid listener judgments from being influenced by voice characteristics
- Recordings were made in a sound-attenuated cabin at the University of Essex
- Data were directly digitized using a Zoom H4 recorder with a sampling rate of 44.1kHz and a resolution of 16Bit

Binary v. Ternary

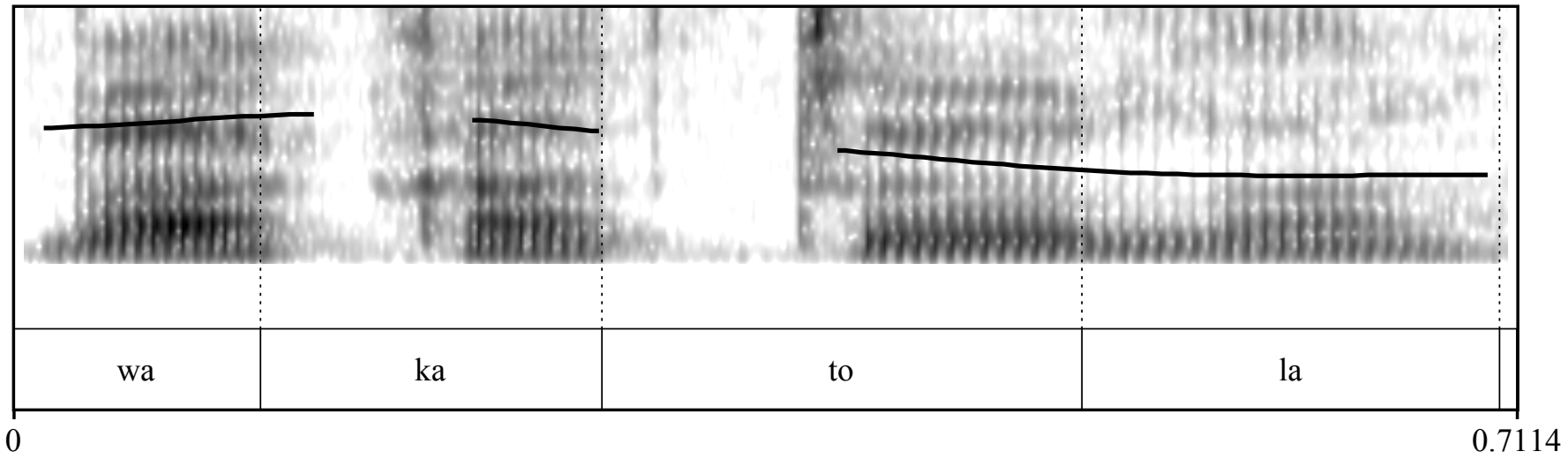


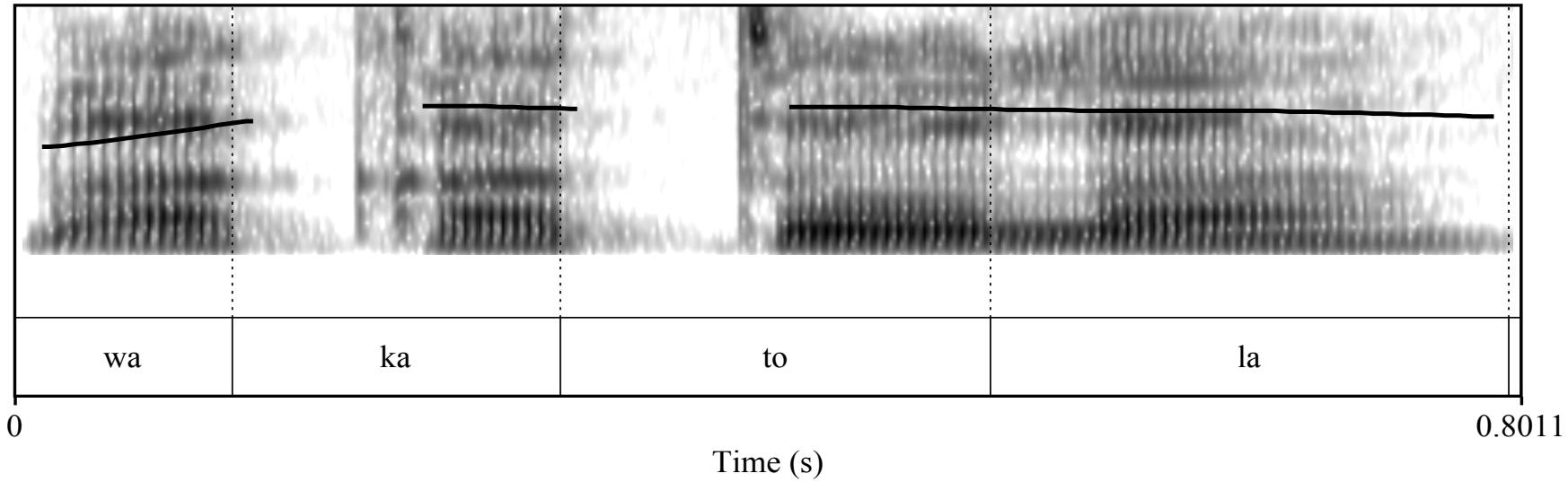


- F0-values of the vowel of the 3rd syllable in the binary-ternary:
- Ternary: average f0-value of 181.4 Hz (SD = 9.0 Hz)
- Binary: average f0-value of 159.8 Hz (SD = 10.7 Hz)

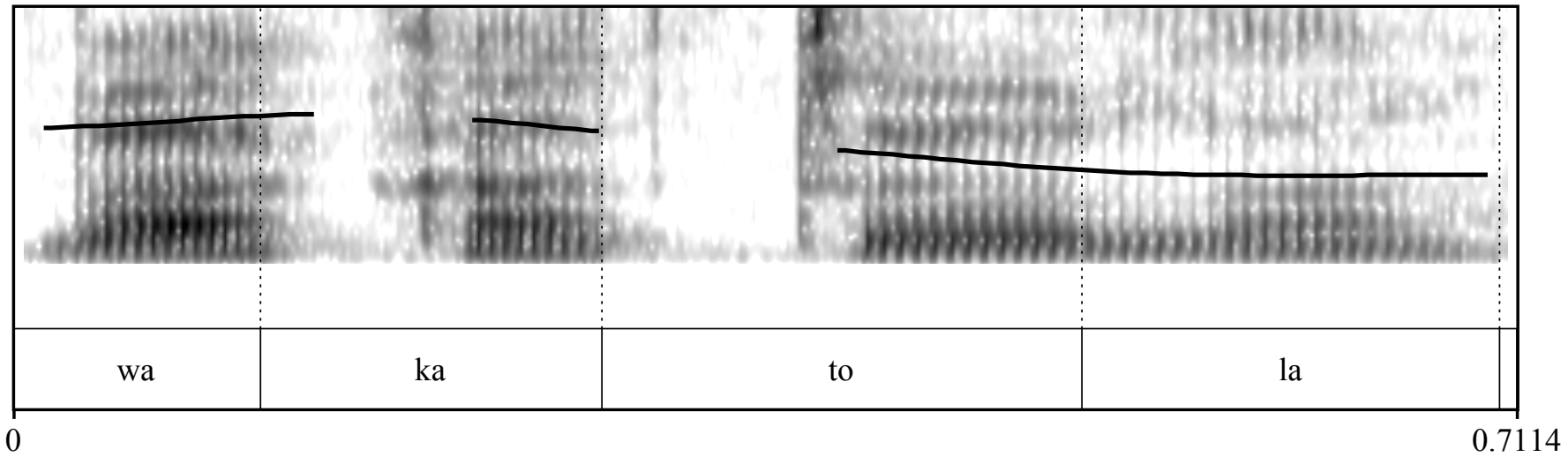
- This difference of 21.6 Hz in average f0-values was significant ($t(31) = 8.6$, $p < 0.0001$)

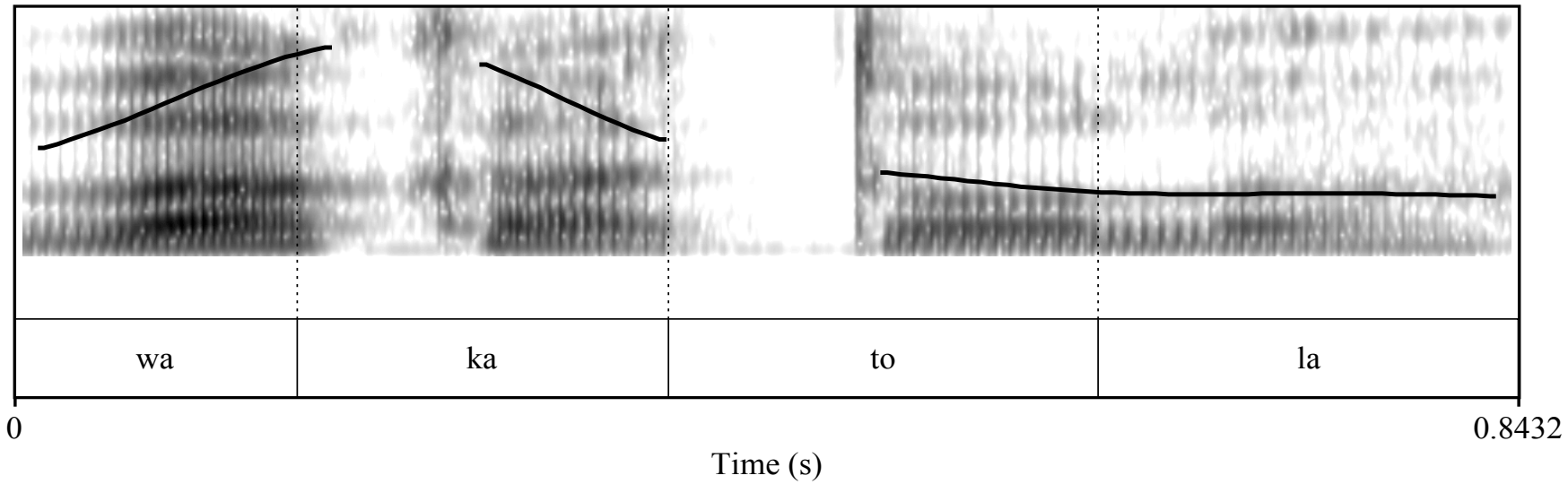
HM v. HH





HM v. HL





- HM-HH and HM-HL: f0-values in the middle of the 3rd syllable vowel:
- HM: average 130.8 Hz (SD = 9.0 Hz)
- HH: average 152.4 Hz (SD = 5.1 Hz)
- HL: average 115.4 Hz (SD = 4.1 Hz)

- Pairwise comparisons showed that these differences in f0-values were significant:
- HM vs. HH: $t(15) = 9.7$, $p < 0.0001$
- HM vs. HL: $t(15) = 6.7$, $p < 0.0001$

- f0-differences were very similar in the three nonce conditions of interest:
- Binary-Ternary v. HM-HH: $t(14) = 0.2$, $p > 0.8$
- Binary-Ternary v. HM-HL: $t(14) = 1.9$. $p > 0.06$

Procedure

- 48 same trials and 48 different trials
- Participants heard each stimulus four times
- Trials were pseudo-randomized to generate the experimental lists: No more than three response types followed each other
- Two stimuli of the same condition were separated by at least one trial of a different condition
- A trial started with a beep and after a period of silence (1000ms) the first stimulus was played (inter-stimulus interval of 2000ms)
- Responses were given by pressing one of two buttons of a button-box
- Responses and response times were recorded for a period of 2000ms

Results

- Responses were converted to d' scores; a measure of sensitivity for a contrast (Macmillan & Creelman, 2005)
- D' scores are based on:
- Hits: correct detection of an existing contrast
- False alarms: detection of a contrast that is not present in the signal
- For the current data set a d' score larger than 1.4 signals a significant sensitivity to the respective contrast

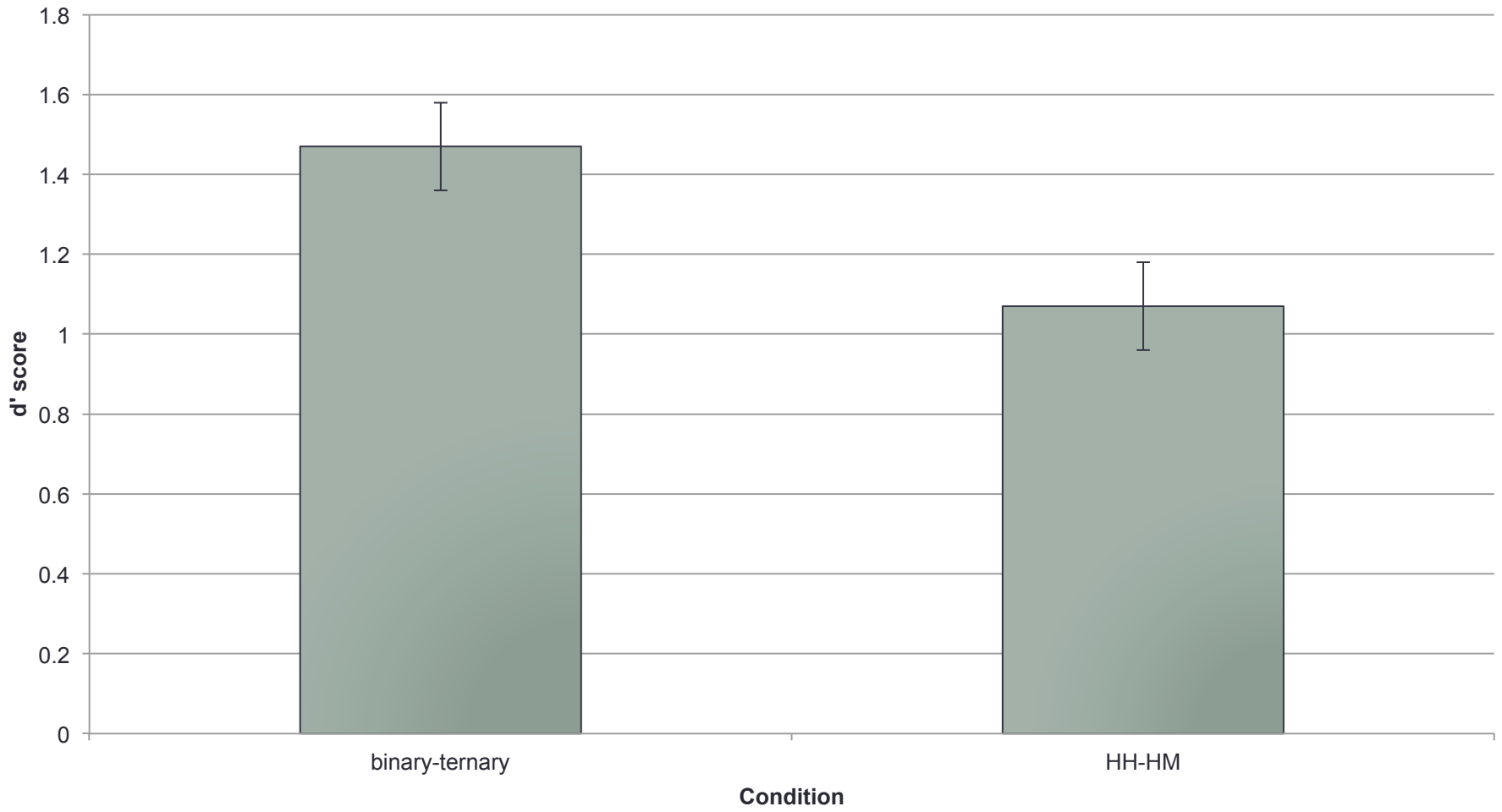
- We first compare participants' sensitivity to the binary vs. ternary spread nonce condition to the HM-HH and HM-HL nonce conditions

- Then we consider the real condition and whether the lexicality of the stimuli in the binary-ternary spread condition affected participants' sensitivity

Bin-Ter v. HM-HH/HL

- Significant main effect of *condition*
- Participants were significantly more sensitive to the distinction between a binary vs. ternary spread (average d' 1.47)
- than to the difference between HM and HH (average d' 1.07, ß = 0.39, SE = 0.11, t = 3.6, $p < 0.001$)
- or to the difference between HM and HL (average d' 1.12, ß = 0.34, SE = 0.09, t = 3.7, $p < 0.001$)
- The difference between HM-HH and HM-HL was not significant ($p > 0.4$)

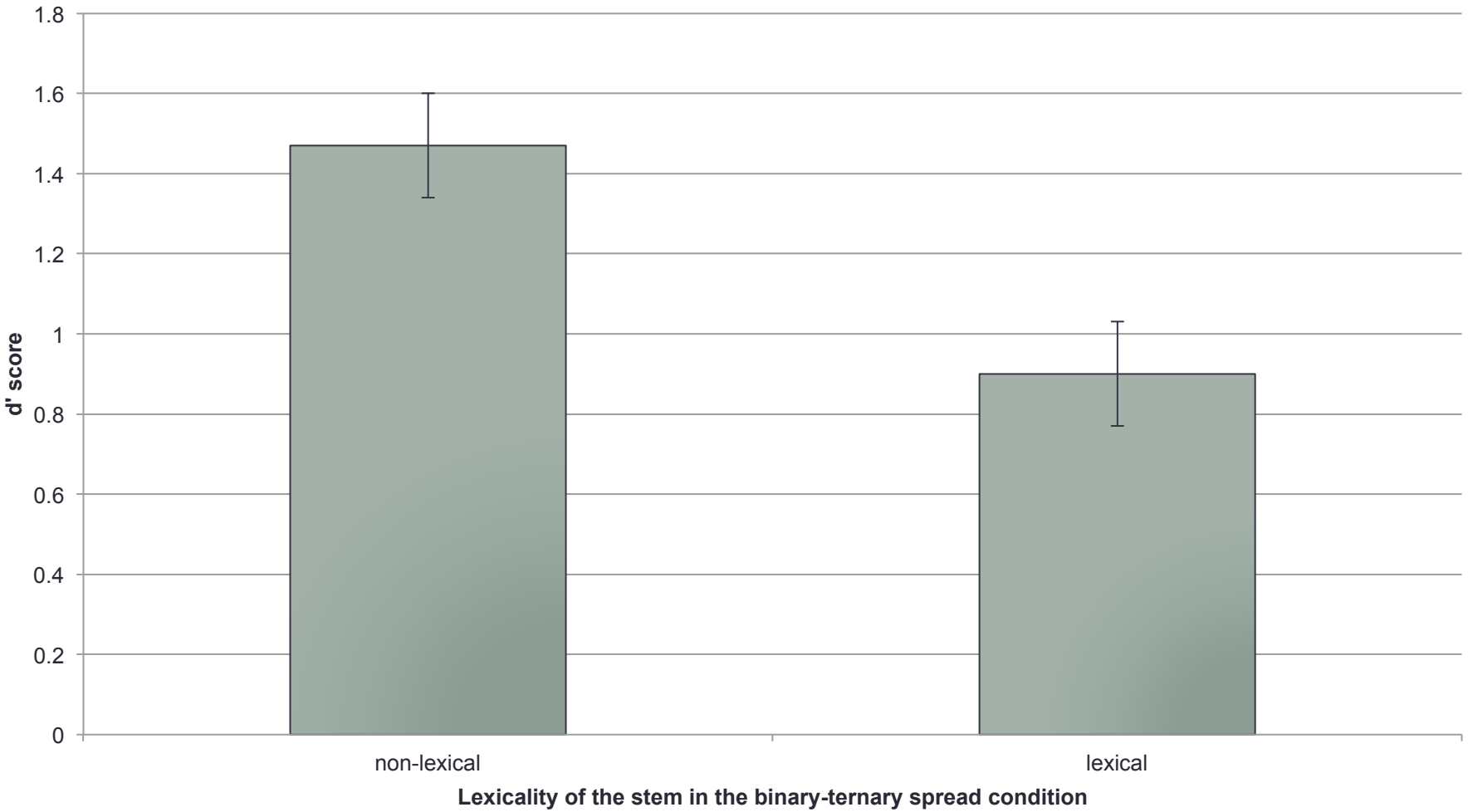
Bin-Ter v. HM-HH/HL



Bin-Ter: real v. nonce

- Results showed a main effect of condition
- Participants were less sensitive to the difference in tone spreading when the stimuli were real words (average d' 0.90)
- than when they were nonce words (average d' 1.47, ß = 0.56, SE = 0.13, t = 4.4, $p < 0.0001$)

Bin-Ter: real v. nonce



Lexical effect

- Participants showed a strong lexical effect: As soon as participants were able to activate a lexical item (based on segmental structure) they paid less attention to tone

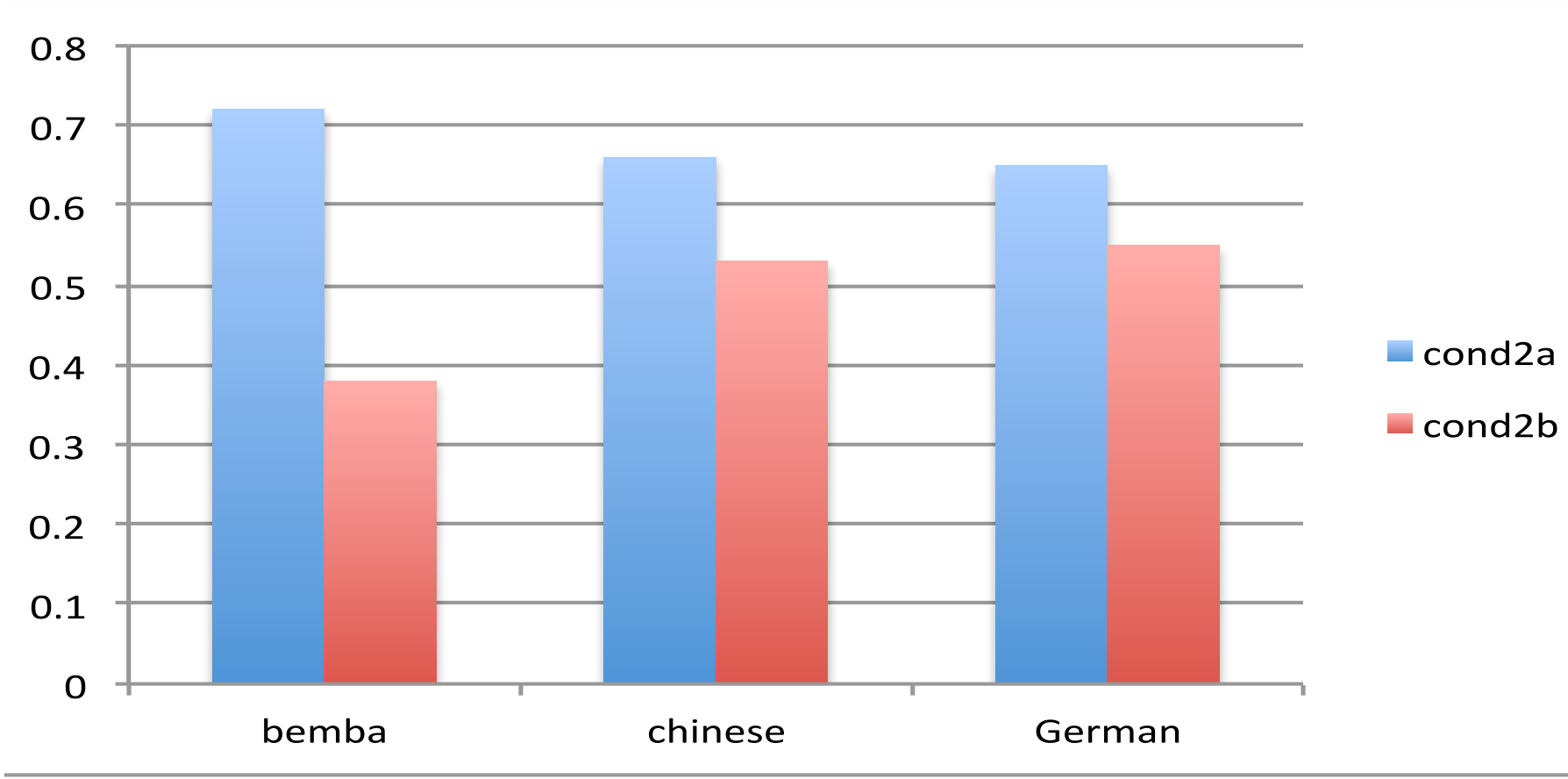
- To establish that there was no other factor that affected the participants performance in real v. nonce words we tested the same conditions with Chinese and German speakers in a follow up Experiment 3

Experiment 3

- Chinese: different kind of tone system (Mandarin speakers)
  - 16 participants (11 female, 5 male)
  - From the University of Essex
  - Aged between 22-27
  - Participated for a small fee

- German: different prosodic system, no tone
  - 14 participants
  - From the University of Konstanz
  - Aged between 21-27

Bemba v. Chinese v. German

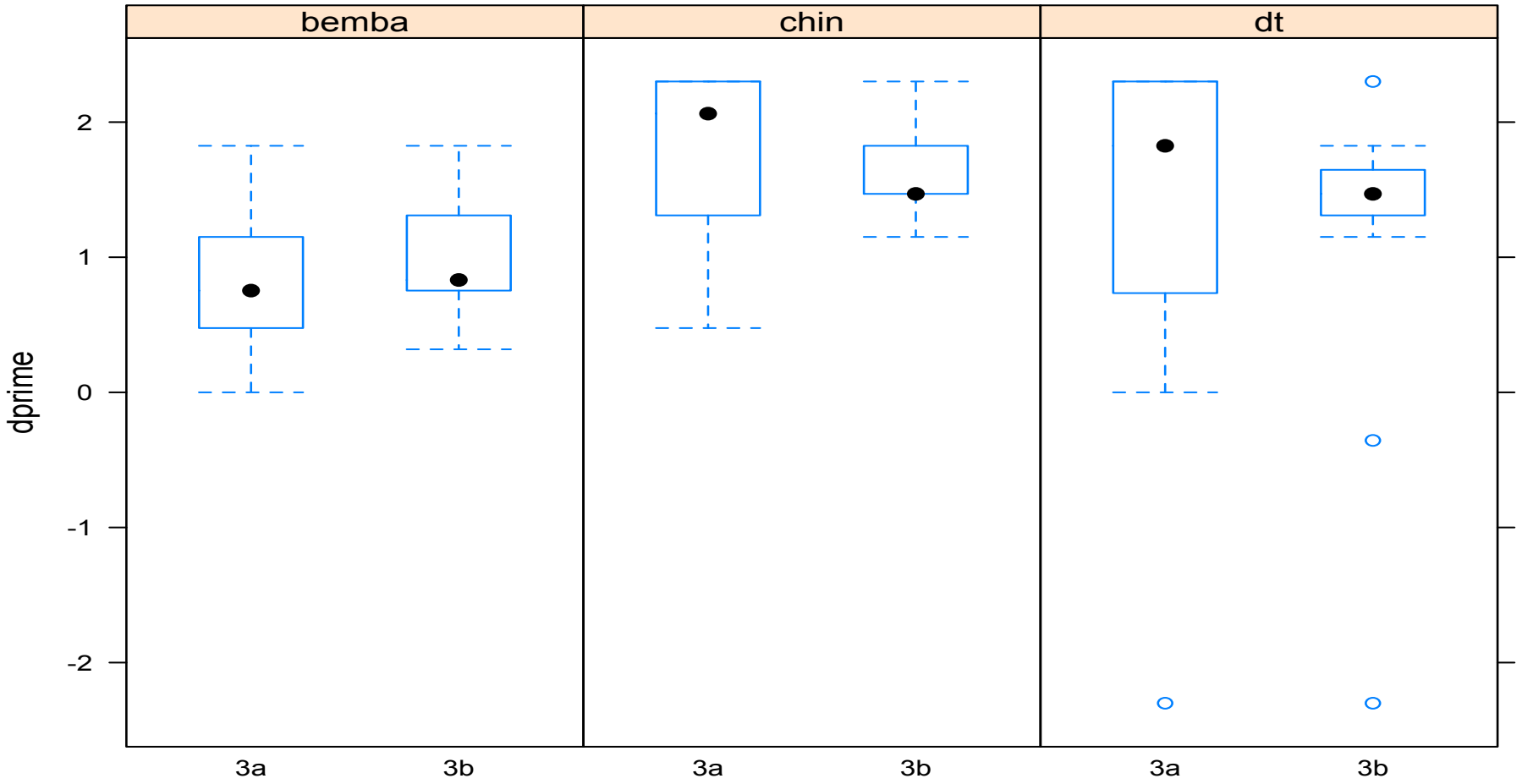


The difference between real (2b) and nonce (2a) words for Chinese and German speakers was not significant ($p > 0.3$)

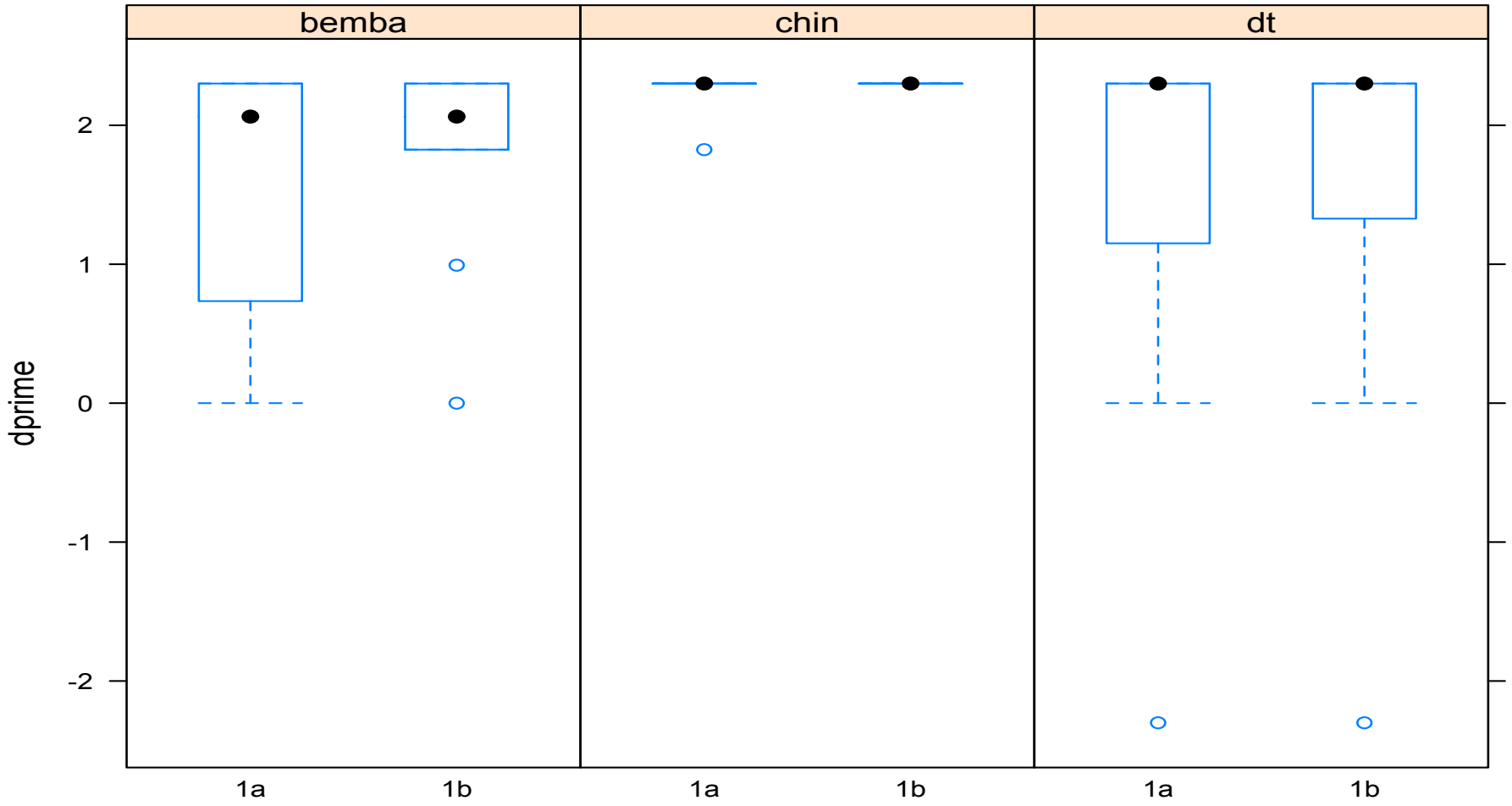
Bemba v. Chinese v. German

- In the nonce condition: no effect of group, all group p-values > 0.3

- In the real condition: group effect; Chinese > Bemba; German > Bemba



D-prime scores split by condition 3a (= HM-HH) vs. 3b (= HM-HL) and language. Main effect of language, Chinese better then Bemba, no effect of condition, no interaction. Germans at chance level.



Conditions 1a and 1b: no main effect, no interaction. (1a: HL-LL , (150-110Hz v. 110-110Hz; 1b: HL-HH, 150-110Hz v. 150-150Hz)

Perception task findings

- Bemba listeners were very sensitive to the difference between binary and ternary spread in the non-lexical condition
- There was a strong lexical effect when real words were used showing that tonal information was overridden by segmental information
- Therefore subjects could not have relied on identifying the two forms as speech from two dialects
- We see that if segmental material triggers lexical items, this blocks tone perception
- This has been attested in studies of stress where higher sensitivity to segmental material has resulted in lexical effects (Peperkamp & Dupoux 2002, Cutler & van Donselaar 2001, Cutler 1986)

- There was significantly better perception of a phonologically relevant contrast (Bin-Ter) than a non-phonologically relevant contrast (HM-HH/HL)

- Ternary spread has a mental representations for speakers of Copperbelt Bemba

Future investigations

- Can we see an incremental lexical effect i.e. gradience in ‘wordiness’ will result in gradience in tone perception?
- We predict that the more real the word gets in terms of segments the less sensitivity to tone there will be.
- Such effects of wordiness have been shown to affect phonemic processing:
- Connine, Titone, Deelman, and Blasko (1997) show that listeners were slowed down more in phoneme-monitoring the less the target resembled a real word (e.g., domestic, tomestic, fomestic).
- In Bantu it must be ensured that the words used do not involve lexically contrastive words or part-words

Conclusions

- Although phonology suggests that ternary spread should not be achieved by one rule, discrimination of nonce words in the binary-ternary condition shows that speakers have a mental representation of ternary domains.
- The validity of ternary spread was supported by the fact that speakers are able to extend their mental knowledge of ternary spread in the production of nonce words.
- In contrast to Asian language tone, Bantu tone is not represented as part of a segment but is autosegmental (Goldsmith 1976).
- Derived tones have a status in phonological processing.

Thank you!

Acknowledgements

- Funding: British Academy & KEDF Research Fund
- Co-researchers: Bettina Braun (& Lee Bickmore)
- Discussion: Aditi Lahiri, audiences at University of Helsinki and UCL
- Speakers and participants (Bemba, Chinese, German)
- Nortec College & Nkrumah University