Phonology **31** (2014). Supplementary materials

Beyond the Iambic-Trochaic Law: the joint influence of duration and intensity on the perception of rhythmic speech

Megan J. Crowhurst
University of Texas at Austin

Amador Teodocio Olivares
Center for the Study and Development of the Indigenous Languages of Oaxaca

Supplementary materials

Appendix A: Stimulus preparation

Experiment 1

Recordings were made using high-quality recording equipment (a Marantz PMD660 solid state digital recorder and a Shure SM10A head-worn unidirectional dynamic close-talking microphone) in a sound-treated room used for recording in the offices of the Center for the Study and Development of the Indigenous Languages of Oaxaca, located near Oaxaca City. Clear tokens of *de* and *ge* were excised from the recorded sequences to be used as precursors, for further manipulation in Praat (Boersma & Weenink 2011). These precursors were selected based on the following criteria: they were closely matched for pitch (the F0 range for *de* was 120–122·5 Hz, and for *ge* 116–121 Hz) and for vowel duration (*de* = 218 ms, *ge* = 216 ms), and care was taken to see that intensity was evenly distributed over the vowel. As noted in §3, consonants in Betaza Zapotec fall into classes described as fortis and lenis. Consistent with the properties of intervocalic lenis stops in Zapotec varieties, the onsets of *de* and *ge* were phonetically voiced and spirantised, [ð] and [ɣ]. Of central importance, as we wished to avoid cues that might lead listeners to interpret either of the syllables as being stressed, both syllables were taken from unemphasised positions in the recorded sequence (consequently, they were softer and shorter than the flanking syllables), and they were closely matched for vowel quality. Since the syllables were coarticulated, F2 rose toward the end of the vowel in [ðe]

and at the beginning of the vowel in [ɣe]. The average values of F1, F2 and F3 over the first half of the vowel in [ðe] and the second half of the vowel in [ɣe] were as follows: *de*: F1: 454·93, F2: 1776·84, F3: 2519·31; *ge*: F1: 460·64, F2: 1825·12, F3: 2493·82. Finally, the precursor syllables had low tone (the default tone) and modal voicing throughout. Each precursor cut from the original recording included the sound wave from the end of its vowel to the end of the preceding vowel (321 ms for *de* and 272 ms for *ge*). The end of a vowel was taken to be the zero-crossing line one cycle after F1 and F2 visibly decreased in intensity. Therefore, each precursor included the spirantised onset, [ð] or [ɣ], and these onsets were matched for intensity and duration. To create base syllables *de* and *ge* for the control sequences (these were further manipulated for the test strings), the following steps were taken: all pitch points for *ge* were raised by 4 Hz to more closely approximate the pitch properties of *de*; each syllable was segmented into onset consonant and vowel; the spirantised onsets were adjusted to 40 ms by adding or removing full cycles; the average intensity of the onsets was scaled to 45 dB; the duration of the vowel in each syllable was adjusted to 255 ms by adding full cycles; the intensity of the vowel was set at 62 dB by changing gain over the full frequency spectrum. Finally, 50 ms of low-intensity (15 dB) noise was added at the midpoint of the [ð] and [ɣ] onsets. This smoothed the transition between syllables of different intensities when they were concatenated into longer sequences. These procedures produced two base syllables (those from which all syllables of different intensities and durations were produced) of almost 350 ms in duration (*de* = 347 ms, *ge* = 344 ms) in which the vowel measured 255 ms in duration and 62 dB in average intensity. Control strings were manufactured by sequencing the base syllables into alternating strings of 10·18 seconds. Two versions were produced, counterbalanced for the string-initial syllable (*de* or *ge*). All sequences ended in the complement of the initial syllable, so that strings consisted of a whole number of paired syllables. In addition to the two control sequences, twelve intensity-varying and twelve duration-varying sequences were produced, as described below, so that in the end, the design called for a total of 26 distinct sequences. String onsets were masked by blending the first five seconds with a segment of white noise. Over this five-second period, the amplitude of the noise was faded out evenly from 62 to 0 dB, and syllables were faded in evenly from 0 dB to their maximum intensities, preserving the alternation of loud and soft syllables. The ends of sequences were backwards masked with a 500 ms, 58 dB segment of white noise that immediately followed the vowel of the final syllable. These masking procedures were also followed for the test sequences used in all of the experiments described here.

Variations in intensity. To create intensity-varying sequences, three variants of the base syllables *de* and *ge* were produced, in which the full-spectrum intensity of the vowel was scaled down in increments of 4 dB by changing overall gain. Sluijter & van Heuven (1996) and Sluijter *et al.* (1997) report that differences between stressed and unstressed syllables in

natural speech in Dutch are more strongly cued by intensity differences concentrated in the frequency range above 500 Hz (so-called spectral tilt), than by differences distributed over the full frequency spectrum (changes in overall gain). This claim is challenged by Ortega-Llebaria *et al*. (2010). In this study, intensity was scaled over the full frequency spectrum, because this method had been used successfully in a similar study, Hay & Diehl (2007). Test strings of 10·18 seconds (the same length as the control strings) were produced by alternating the 62 dB *de* syllable with a softer *ge* syllable (58, 54 or 50 dB). Six test sequences were used, three with loud *de* and a counterbalanced set of three with (softer) *ge* as the string-initial syllable. A counterbalanced set of sequences was produced by exchanging the parameter settings for *de* and *ge* (so that *ge* was the louder and *de* the softer syllable). In all, there were twelve intensity-varying sequences.

Variations in duration. For the duration-varying sequences, three variants of *ge* shorter than the base syllable in increments of 41 ms were produced by adding or removing full voicing periods from the steady-state portion of the precursor vowels. The increments used for duration were larger than those used by Hay & Diehl (2007), and smaller than those used by Bhatara *et al*. (2013) (25 and 50 ms in the two studies respectively). The average intensities of the shortened vowels were readjusted to 62 dB to match the intensity of the base syllables. A set of test strings was created by alternating the 255 ms *de* syllable with a shorter variant of *ge* (214, 173 and 132 ms in duration). Six test sequences were used, three with (long) *de* and three with (short) *ge* as the string initial syllable. Sequences were made to be as close to 10 seconds as possible without truncating syllables or pairs of syllables (the range was 9·60 to 10·73 seconds). A counterbalanced set of sequences, in which *ge* was the longer and *de* the shorter syllable, was produced by exchanging the parameter settings for *de* and *ge*. These procedures produced twelve duration-varying sequences in all.

Experiment 2

The control sequences prepared for Experiment 1 were used again in Experiment 2. Test sequences were produced by alternating the syllables created for Experiment 1 in new combinations according to the design in Table X (App. C). The magnitude scales for intensity and duration introduced in Fig. 1 were also critical in Experiment 2, and the same increments of variation for intensity and duration were used, as described earlier. Because this was a fully counterbalanced design, the set of test sequences in Experiment 2 was much larger than for Experiment 1. The design (shown in Table X) called for different levels on the MOI and MOD scales to be crossed, producing 74 (72 test and two control) sequences in all. These were divided into two balanced sets, Set 1 and Set 2. For the Cooperating condition, the softest and longest (50 dB, 255 ms) token of *de* was alternated with louder and shorter *ge* tokens. The three intensity levels used were 62, 58 and 54 dB, and the three duration levels were 214, 173 and 132 ms. As before, the full set of 18 sequences produced in this manner was counter-

balanced as to whether the sequence began with the longer or shorter syllable. In addition to the 36 test sequences produced in this way, Set 1 included the two counterbalanced control sequences (a total of 38 sequences). Set 2 consisted of the two control sequences and 36 sequences in which the parameter settings for *de* and *ge* were exchanged (so that *de* was now the longer softer syllable).

Appendix B: Instructions, printed materials and experimental procedures

Experiment 1

Print materials. Each participant was given a stapled booklet containing six pages. The first was a demonstration page with instructions in Zapotec and lines for four practice trials. The remaining pages were response sheets, one for each block. These were printed with 28 numbered lines, one for each trial in the block. As many Betazans have had limited schooling, a circle and a square were used as placeholders on numbered lines to minimise text on the page. The shapes on alternate lines were shaded differently to provide contrast between lines.Two response columns on the left were labelled 'GE-DE' and 'DE-GE'. (The order of these columns was not counterbalanced; 'GE-DE' was always leftmost, so as not to complicate instructions.) Two additional columns on the right containing the digits '1' and '2', and labelled *lekzena'* 'certain' and *dzekzelan* 'less sure'. Participants were asked to circle one of the digits to indicate confidence ratings.

Experimental procedure. Participants in Experiment 1 were tested in two groups in rooms in a private home ($n = 6$) and a municipal building ($n = 26$) in San Melchor de Betaza. Stimuli were presented in six test blocks, which were preceded by four practice trials. In each block, the twelve intensity-varying, twelve duration-varying and two iterations of each control string (a total of 28 sequences) were intermingled. Sequence order within blocks was randomised by the SuperLab program each time a block was run. Test blocks were preceded by four practice trials, and participants received feedback following these before moving on. Feedback and assistance were provided to those who needed it by the second author and (in the larger session) by four school teachers who were on hand for the duration of the experiment, and who received the same compensation as participants. In explaining the task, participants were instructed to wait until after the noise had faded completely and the syllables had come up to their full volume before making a decision, and to respond while the recording was still playing. The experiment was controlled by the software program SuperLab 4.0 (Cedrus Corporation) running on a MacBook Pro laptop computer. The software program was managed by a member of the community who had been trained for this purpose. Sequences were presented in free field at full volume over a portable Bose Superlink speaker positioned in front of the group, connected by cable to the computer's audio output jack.

The Zapotec language environment was reinforced in the following ways: informed consent procedures, instructions and any discussion prior to testing were accomplished verbally in Zapotec by the second author, who conducted the experiment. For the larger group, the verbal instructions provided by the experimenter were reinforced by diagrams drawn with markers on large sheets of paper, taped to a wall at the front of the room. During the experiment, the second author kept pace by calling out the number of each trial in Zapotec. The first author, an outsider in this context, was not in the room during test sessions. An information sheet written in Spanish, and approved by the Institutional Review Board of the University of Texas at Austin was given to participants at the end of each session. The entire experiment took approximately 50 minutes to complete.

Experiments 2–4

Response sheets consisted of numbered blanks in which participants were instructed to write 'd' for 'dege' or 'g' for 'gede'. In the first experiment, we used two columns headed GE-DE and DE-GE, whose order was not counterbalanced. The method used in Experiments 2–4 was changed to control for the possibility that column order might influence listeners' decisions.

Appendix C: Response data

		magnitude						
		−3	−2	−1	0	1	2	3
intensity	*dege*	150 (44)	185 (55)	179 (53)	314 (46)	166 (49)	215 (65)	160 (48)
	gede	188 (56)	150 (45)	157 (47)	363 (54)	171 (51)	118 (35)	173 (52)
duration	*dege*	168 (50)	126 (38)	186 (56)	*as*	172 (51)	138 (41)	174 (52)
	gede	168 (50)	204 (62)	147 (44)	*above*	166 (49)	195 (59)	162 (48)

Table X

Design and response data for Experiment 1, showing number and proportion (as a percentage) of *dege* and *gede* responses for each cell. Magnitude level 0 represents the control condition.

			cooperating (*de* = long/soft)			competing (*de* = long/loud)		
duration of *de*	3	*dege*	14 (13)	13 (12)	25 (23)	76 (71)	74 (69)	100 (93)
		gede	92 (87)	95 (88)	83 (77)	31 (29)	33 (31)	8 (7)
	2	*dege*	7 (6)	16 (15)	24 (22)	77 (71)	93 (87)	88 (81)
		gede	101 (94)	91 (85)	84 (78)	31 (29)	14 (13)	20 (19)
	1	*dege*	12 (11)	14 (13)	13 (14)	85 (79)	92 (86)	103 (95)
		gede	95 (89)	94 (87)	83 (86)	22 (21)	15 (14)	5 (5)
			−3	−2	−1	1	2	3
			intensity of *de*					
	−1	*dege*	9 (8)	5 (5)	17 (16)	86 (81)	101 (94)	106 (98)
		gede	98 (92)	102 (95)	91 (84)	20 (19)	7 (6)	2 (2)
	−2	*dege*	13 (13)	8 (7)	26 (24)	93 (86)	100 (93)	105 (97)
		gede	91 (87)	99 (93)	82 (76)	15 (14)	8 (7)	3 (3)
	−3	*dege*	8 (7)	15 (14)	26 (25)	97 (90)	102 (95)	102 (95)
		gede	99 (93)	91 (86)	80 (75)	11 (10)	5 (5)	5 (5)
			competing (*de* = short/soft)			cooperating (*de* = short/loud)		

Table XI

Design and response data for Experiment 2, showing number and proportion (as a percentage) of *dege* and *gede* responses for each cell. Results for the control sequence (magnitude 0; not shown); *dege* = 87 (40%), *gede* = 129 (60%).

			cooperating			duration	competing		
magnitude of duration (MOD)	4	*dege*	25 (25)	27 (27)	36 (36)	51 (50)	65 (65)	70 (74)	75 (75)
		gede	75 (75)	73 (73)	65 (65)	50 (50)	35 (35)	25 (26)	25 (25)
	3	*dege*	26 (26)	24 (24)	30 (29)	47 (46)	67 (66)	71 (71)	81 (79)
		gede	75 (74)	75 (76)	72 (71)	55 (54)	34 (34)	29 (29)	21 (21)
	2	*dege*	15 (15)	21 (21)	19 (19)	46 (45)	79 (78)	73 (73)	90 (88)
		gede	86 (85)	80 (79)	82 (81)	56 (55)	22 (22)	27 (27)	12 (12)
	1	*dege*	14 (14)	13 (13)	14 (14)	40 (40)	79 (77)	80 (80)	88 (88)
		gede	87 (86)	88 (87)	87 (86)	60 (60)	23 (23)	20 (20)	12 (12)
	0	*dege*		intensity		36 (36)	80 (79)	88 (87)	77 (76)
		gede				64 (64)	21 (21)	13 (13)	24 (24)
			−3	−2	−1	0	1	2	3
			magnitude of intensity (MOI)						

Table XII

Design and response data for Experiment 3, showing number and proportion (as a percentage) of *dege* and *gede* responses for each cell. Magnitude of intensity reflects the intensity of *de* relative to *ge*, and magnitude of duration reflects increasing duration of *de* relative to *ge*.

			cooperating			duration	competing		
magnitude of duration (MOD)	4	*dege*	68 (38)	51 (29)	53 (30)	64 (36)	77 (43)	124 (69)	84 (47)
		gede	111 (62)	127 (71)	123 (70)	114 (64)	103 (57)	56 (31)	95 (53)
	3	*dege*	60 (33)	44 (25)	59 (33)	61 (34)	90 (50)	103 (57)	96 (53)
		gede	120 (67)	135 (75)	119 (67)	119 (66)	90 (50)	77 (43)	84 (47)
	2	*dege*	68 (38)	57 (32)	45 (25)	81 (45)	99 (55)	114 (64)	71 (40)
		gede	110 (62)	122 (68)	134 (75)	99 (55)	80 (45)	64 (36)	108 (60)
	1	*dege*	50 (28)	53 (30)	48 (27)	78 (44)	105 (59)	120 (67)	117 (65)
		gede	130 (72)	124 (70)	128 (73)	100 (56)	74 (41)	60 (33)	62 (35)
	0	*dege*		intensity		165 (46)	115 (65)	123 (69)	127 (71)
		gede				194 (54)	63 (35)	56 (31)	51 (29)
			−3	−2	−1	0	1	2	3
			magnitude of intensity (MOI)						

Table XIII

Design and response data for Experiment 4, expressed as number and proportion (as a percentage) of *dege* and *gede* responses for each cell (cf. Table XII).