El Aissati, McQueen & Cutler:

Finding words in a language that allows words without vowels

SUPPLEMENTARY MATERIAL

1. Schwa condition, Tarifiyt experiment

1.1 *Justification*

In the Berber literature, some authors claim that schwa is realised in syllables transcribed and perceived as vowelless (see, e.g., Coleman, 1999). If this is not perceptually relevant, then contexts with schwa could be perceived as vowelless and pattern like consonant contexts, differently from contexts with full vowels. However, in other PWC studies, schwa contexts have always patterned with contexts with full vowels (e.g., McQueen & Cutler, 1998; Norris et al., 2001). A standard PWC effect in Berber would therefore predict that contexts with schwa would pattern like contexts with full vowels, differently from consonant contexts.

1.2 *Materials*

Each context also occurred in a version with the reduced vowel schwa (e.g., for *fad*, the context *egh* [əɣ]) as well as the consonant and syllable contexts reported in the main text.

1.3 *Results*

Mean RT in the schwa context condition was 656 ms, mean error rate 27%. Neither RTs nor errors differed from the other two conditions.

1.4 *Conclusion*

The contexts with schwa patterned like contexts with full vowels, and also like consonant contexts. Since these also did not differ, the contexts with schwa do not add to the interpretation of the results.

1. 5 *References*

Coleman, J.S. (1999). The nature of vocoids associated with syllabic consonants in Tashlhiyt Berber. In J.J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, & A.C. Bailey (Eds.), *Proceedings of the XIVth International Congress of Phonetic Sciences* (pp. 735-738). San Francisco: American Institute of Physics, University of California Press.

McQueen, J.M., & Cutler, A. (1998). Spotting (different types of) words in (different types of) context. In R.A. Mannell & J. Robert-Ribes (Eds.), *Proceedings of the 5th International Conference on Spoken Language Processing* (Vol. 6, pp. 2791-2794). Sydney: Australian Speech Science and Technology Association.

Norris, D., McQueen, J.M., Cutler, A., Butterfield, S., & Kearns, R. (2001). Language-universal constraints on speech segmentation. *Language and Cognitive Processes*, **16**, 637-660.

2. Materials, Tarifiyt Experiment

Target	Translation	Consonant Context	Reduced-syllable Context	Full-syllable Context
mun	to accompany	ħmun	əħmun	aħmun
batˤəl	free of charge	ħbatˤəl	əħbatˤəl	uħbatˤəl
niy	to mount	qniy	əqniy	uqniy
siwəl	to speak	ɣsiwəl	əɣsiwəl	uɣsiwəl
tadunt	fat	ɣtadunt	əɣtadunt	aɣtadunt
dəwwəx	to get dizzy	ʒdəwwəx	əʒdəwwəx	uʒdəwwəx
səqsa	to ask	ʒsəqsa	əʒsəqsa	aʒsəqsa
tirjin	coal	ztirjin	əztirjin	aztirjin
laɣa	to call	zlaɣa	əzlaɣa	azlaɣa
badu	edge	ʕbadu	əʕbadu	aʕbadu
qabəl	to face so/sth	ʕqabəl	əʕqabəl	uʕqabəl
tala	source of water	btala	əbtala	ubtala
ɣayitˤa	cookies	bɣayitˤa	əbɣayitˤa	abɣayitˤa
middən	people	ħmiddən	əħmiddən	uħmiddən
qawit	peanuts	ħqawit	əħqawit	aħqawit
marməd	to abuse verbally	qmarməd	əqmarməd	aqmarməd
sird	to wash	qsird	əqsird	uqsird
fad	thirst	ɣfad	əɣfad	aɣfad
tˤumatˤic	tomatoes	ɣtˤumatˤic	əɣtˤumatic	uɣtˤumatˤic
limart	a sign	ʒlimart	əʒlimart	uʒlimart
xali	maternal uncle	zxali	əzxali	uzxali
tamar	to suffer	ztamar	əztamar	aztamar
gəwwəd	to guide	ʕgəwwəd	əʕgəwwəd	aʕgəwwəd
tanəwwart	a flower	ʕtanəwwart	əʕtanəwwart	uʕtanəwwart
lajar	grace	blaʒar	əblaʒar	ablaʒar
ʕarn	to push	bʕarn	əbʕarn	ubʕarn
filu	thread	ħfilu	əħfilu	aħfilu
gəʕʕəd	to climb up	ħgəʕʕəd	əħgəʕʕəd	uħgəʕəd
batˤatˤa	potatoes	qbatˤatˤa	əqbatˤatˤa	aqbatˤatˤa
ʕawəd	to repeat	qʕawəd	əqʕawəd	uqʕawəd
saməħ	to pardon	ɣsaməħ	əʕsaməħ	uʕsaməħ
tazart	figs	ɣtazart	əɣtazart	aɣtazart
gənfa	to recover	jgənfa	əjgənfa	ajgənfa
ʕumm	to swim	jʕumm	əjʕumm	ajʕumm
qibalt	in front of	zqibalt	əzqibalt	azqibalt
ɣana	desire	zɣana	əzɣana	uɣhana
ru	to cry	ʕru	əʕru	aʕru
ʒawar	to be pregnant	bʒawar	əbʒawar	ubʒawar
ʕam	one year	bʕam	əbʕam	abʕam

3. Materials, Tashelhiyt Experiment

Target	Translation	Consonant Context	Reduced-syllable Context	Full-syllable Context
fad	thirst	qfad	əqfad	aqfad
middən	people	ħmiddən	əħmiddən	aħmiddən
taɣawsa	thing	xtaɣawsa	əxtaɣawsa	axtaɣawsa
kabab	kebab	ʒkabab	əʒkabab	aʒkabab
tirgi	burning coal	ztirgi	əztirgi	aztirgi
duf	to guard	zduf	əzduf	azduf
tigəmmi	house	qtigəmmi	əqtigəmmi	uqtigəmmi
siyyəs	to boil	ħsiyyəs	əħsiyyəs	uħsiyyəs
tasa	liver	ħtasa	əħtasa	uħtasa
nu	to be ripe	xnu	əxnu	uxnu
fərtəttu	bat	jfərtəttu	əjfərtəttu	ujfərtəttu
bukəd	to go blind	zbukəd	əzbukəd	uzbukəd
tifawt	light	qtifawt	əqtifawt	aqtifawt
bidd	to stand up	ħbidd	əħbidd	aħbidd
malɣəf	brains	xmalɣəf	əxmalɣəf	axmalɣəf
sawəl	to speak	xsawəl	əxsawəl	axsawəl
bahra	plenty	ʒbahra	əʒbahra	aʒbahra
matiʃa	tomatoes	zmatiʃa	əzmatiʃa	azmatiʃa
mun	to accompany	ħmun	əħmun	uħmun
saqsa	to ask	ħsaqsa	əħsaqsa	uħsaqsa
saməħ	to pardon	xsaməħ	əxsaməħ	uxsaməħ
tawəryit	rope	ʒtawəryit	əʒtawəryit	uʒtawəryit
rar	to back	ʒrar	əʒrar	uʒrar
matˁəl	to be late	zmatˁəl	əzmatˁəl	uzmatˁəl
sərs	to put down	qsərs	əqsərs	aqsərs
liqamt	mint	ħliqamt	əħliqamt	aħliqamt
su	to drink	xsu	əxsu	axsu
timəzgyida	mosque	xtiməzgyida	əxtiməzgyida	axtiməzgyida
dawa	to cure	ʒdawa	əʒdawa	aʒdawa
lal	owner	zlal	əzlal	azlal
fitəllis	moth	qfitəllis	əqfitəllis	uqfitəllis
kuʃəm	to be crippled	ħkuʃəm	əħkuʃəm	uħkuʃəm
gawər	to sit down	ħgawər	əħgawər	uħgawər
tadˁsa	laughter	xtadˁsa	əxtadˁsa	uxtadˁsa
tadunt	fat	ʒtadunt	əʒtadunt	uʒtadunt
batˁatˁa	potatoes	zbatˁatˁa	əzbatˁatˁa	uzbatˁatˁa

4. Tashelhiyt Experiment

4.1 Method

4.1.1 Participants

Fifty five student volunteers (17 female, mean age 21 years, range 18-23 years), from Agadir and the surrounding area in Morocco, were paid to take part. All were native speakers of Tashelhiyt Berber, with no known hearing problems.

4.1.2 Materials

Forty eight target Tashelhiyt words were selected. These words varied in length between one and four syllables, and all began with a Consonant-Vowel (CV) sequence. Nonsense sequences for each target word were made by adding three contexts onto the beginning of each word. Preceding single-consonant (C), reduced-syllable (schwaC, here transcribed eC) and full-syllable (VC) contexts were selected. Five consonants were chosen (/q/, /ħ/, /x/, /ʒ/, and /z/, transcribed here as *q, h, x, j*, and *z* respectively). Each word was paired with one of these consonants, which served as the single-consonant context for that word, and as the consonant in the syllable contexts (e.g., for the target *fad,* 'thirst', the three target-bearing nonsense sequences were *qfad, eqfad*, and *aqfad*). The sequences of consonants straddling target-word onsets (e.g., *qf* for *fad*) all occur word-internally in Tashelhiyt, and thus do not signal phonotactically mandatory word boundaries.

An additional 78 stimuli were made. Sixty were fillers, that is, they were nonsense sequences which did not contain real Tashelhiyt words. They were made by analogy to the target-bearing items: CV-initial nonwords were preceded by a single consonant, a VC syllable with a reduced vowel, or a VC syllable with a full vowel (there were 20 of each of these three context types). The remaining 18 items were used in the practice block. Six contained targets and twelve did not (within each of these sets, one third began with each of the three context types).

4.1.3 Stimulus preparation

The stimuli were recorded by a male native speaker of Tashelhiyt Berber who was unaware of the experiment's purpose. The recording was made in a sound-attenuated

booth onto Digital Audio Tape, sampling at 48 kHz. Targets in full-syllable contexts were recorded with both /a/ and /u/ in the context, and multiple recordings were made of each item. All materials were transferred to computer (downsampling to 16 kHz, 16 bits) and were selected, measured and manipulated using the Xwaves speech editor.

The target-bearing items were made by cross-splicing in the same way as for the Tarifiyt experiment. In some cases it was not possible to create usable triplets of target-bearing items, either because there were no suitable recordings of a target and/or a context (e.g., there was no detectable schwa, vowels were not spoken correctly, or an epenthetic schwa was inserted between the consonant of the context and the initial consonant of the target), or because cross-splicing created audible discontinuity. A final set of 36 targets and their contexts was selected in which the targets and contexts were spoken as intended, and in which the splicing was not detectable. In half of the final items the vowel in the full-syllable context was /a/; in the other half it was /u/. Fillers and practice items were not spliced; a clearly-articulated token of each of these was selected.

4.1.4 Design and procedure

Three counter-balanced versions of the experiment were constructed. Each version consisted of the same practice block followed by an experimental block that contained all 36 targets and all 60 fillers. What differed across the three experimental blocks were the contexts in which the targets appeared. Context type was counterbalanced over versions, such that each version contained 12 targets in each of the three contexts (single consonant, reduced syllable and full syllable contexts) and which contained each target only once, but such that in the experiment as a whole all targets appeared in all contexts. Target-bearing and filler items were mixed in a quasi-random order such that there was always at least one filler item between any two target-bearing items. The running order of targets and fillers was the same in all three versions; the only difference across versions lay in which context was used for a given target.

Participants were tested individually in a quiet room. They were randomly assigned to one of the three versions of the experiment. Auditory stimuli were presented to them over headphones. The participants were told that they would hear a list of nonsense sequences and that their task was to try to spot any real Tashelhiyt words embedded at the end of the nonsense sequences. They were asked to press a response

button (with a finger of their dominant hand) as quickly as possible if they thought they had spotted a real word, and then to say that word aloud. The oral responses of the participants were recorded. Stimulus presentation (Inter-Stimulus Interval: 3.5 sec.) and logging of response latencies were controlled by NESU software on a laptop computer.

Reaction Times (RTs) were measured from stimulus onset, but adjusted prior to data analysis so as to measure from target offset (by subtracting the total duration of the appropriate stimulus from each raw RT). Participants' oral responses were examined. All manual responses which were accompanied by either an incorrect oral response (i.e., a word other than the intended target) or no oral response were treated as errors (this occurred on 4.4% of all experimental trials).

4.2 Results

The data from six participants were not analysed, either because their oral and/or manual results were incompletely recorded due to technical problems (three subjects), or because they gave Arabic oral responses (one subject), or because they were excluded because they missed at least two thirds of all targets (two subjects). The data from 11 target words (*duf, siyyes, nu, malʏef, mun, taweryit, rar, su, dawa, lal* and *fitellis*) were also excluded because, collapsing over the three context conditions, these items were missed by more than 60% of the participants. Mean RTs for the remaining 49 participants and 25 words in each of the three context conditions are given in Table 1. Note that the pattern of results was the same when no participants or items were excluded.

Table 1. Experiment 1: Mean Reaction Times (RTs, in milliseconds from target word offset), and Mean Percentage Error Rates

	Consonant Context	**Reduced-vowel Context**	**Full-vowel Context**
RT	546	574	663
Error	13%	14%	16%

Targets in full-syllable contexts were detected much more slowly than in either reduced-syllable or consonant contexts. There was little difference in latency between the latter two conditions, and no substantial differences among all three conditions in error rates. Repeated-measure Analyses of Variance (ANOVAs) were performed by participant (F1) and item (F2) on the latency and error data. There was an effect of the factor context in the latency analysis ($F1(2,92) = 9.22$, $p<.001$; $F2(2,48) = 4.86$, $p=.01$) but not in the error analysis ($F1(2,92) = 1.32$, $p>.2$; $F2(2,48) = 1.42$, $p>.2$). Planned comparisons showed that the context effect in RTs was due primarily to the slower responses in the full-syllable contexts (consonant vs. reduced syllable, 28 ms, on average: $F1(1,46) = 1.65$, $p>.2$, $F2(1,24) = 1.23$, $p>.2$; consonant vs. full syllable, 117 ms, on average: $F1(1,46) = 16.49$, $p<.001$, $F2(1,24) = 7.69$, $p=.01$; reduced vs. full syllable, 89 ms, on average: $F1(1,46) = 7.14$, $p=.01$, $F2(1,24) = 4.17$, $p=.05$).

An analysis of target length (number of syllables) compared responses to nine monosyllabic targets (two could not be included in the RT analysis because there were no responses to those targets in at least one context condition) to those to 19 bisyllabic targets (one was excluded for a similar reason in the RT analysis); words with three and four syllables were excluded (seven words and one word, respectively). Tashelhiyt participants were slower (by 151 ms, on average; $F1(1,46) = 36.95$, $p <.001$; $F2(1,23) = 10.70$, $p<.005$) and less accurate (by 36%, on average; $F1(1,46) = 218.00$, $p<.001$; $F2(1,26) = 9.04$, $p<.01$) in spotting monosyllabic than bisyllabic words.

4.3 Conclusion

The results of this experiment suggest that the Possible Word Constraint plays no role in speech segmentation in Tashelhiyt.