FRENCH LISTENERS' DEAFNESS TO TASHLHIYT BERBER /bi/-/bbi/

P. Hallé[a,b] *& R. Ridouane*[a]

[a]LPP (CNRS-Paris 3), France; [b]LPNCog (CNRS-Paris 5), France
pierre.halle@parisdescartes.fr; rachid.ridouane@wanadoo.fr

ABSTRACT

In a discrimination experiment on Tashlhiyt Berber singleton-geminate contrasts (i.e., duration contrasts), we find that French listeners are more sensitive to silent closure duration in word-final /t/ than to voiced murmur, or even, frication duration in initial position. Native listeners of Tashlhiyt perform near-ceiling on all these contrasts. We propose that native listeners of French, in which gemination is not phonemic, have not acquired quantity contrasts but yet retain a perhaps universal sensitivity to rhythm, or more specifically to inter-gestural timing.

Keywords: nonnative speech perception, Tashlhiyt Berber, French, geminate obstruents

1. INTRODUCTION

The seminal studies by Abramson and Lisker [2, 3] showed how a VOT continuum was perceived into different categories and with different categorical boundaries across listener's native languages. Spanish-speaking listeners [3], for example, would segment the continuum into two VOT categories: prevoiced, and voiceless, in agreement with their production of the Spanish contrast. French listeners perform very similarly [9] presumably because French uses the same phonetic settings as Spanish to distinguish its two phonemic voicing categories in stops. On these grounds, French listeners should not be able to discriminate prevoiced stops that differ in prevoicing duration. Prevoicing duration is considered as phonologically distinctive in those languages that contrast singleton and geminate prevoiced stops in all positions, such as Pattani Malay [1] and Tashlhiyt Berber [8], or in word-medial position only, such as Italian, (The former languages are few.) By this criterion, prevoicing duration, that is, voiced stop quantity, is not contrastive in French. Yet, geminates do occur in French at word boundaries (*avec quoi* /avɛkkwa/ 'small size'), or word-internally following schwa deletions (*netteté* /nɛtte/ 'neatness'), and quite often in the future or conditional, as opposed to the imperfect tense (*courais*$_{1sg.imp}$ vs. *courrais*$_{1sg.cond}$ /kurɛ/ vs. /kurrɛ/ 'to run'). In particular, geminate voiced stops at word boundary may contrast with a singleton counterpart, as in *là dedans* vs. *la dent*, /laddɑ̃/ vs. /ladɑ̃/, 'in there' vs. 'the tooth'. As we noted, such minimal pairs are not phonemic from the usual phoneme repertoire perspective. This situation raises two questions: (i) Can French listeners distinguish such minimal pairs? (ii) If they can, what is the phonetic basis of their discrimination capacity?

With respect to the first question, we may surmise that French listeners distinguish *là dedans* from *la dent* within a sentence context, that is, given sufficient top-down information. Studies testing the comprehension of such utterances presented in isolation are few. Meisenburg's data [7] suggest that listeners do distinguish *il l'a dit* from *il a dit* (/illadi/-/iladi/) in production and perception but distinguish less well other pairs. As for the second issue, a number of phonetic characteristics are possible cues to the singleton-geminate distinction and could be used in perception. The major cue, however, logically should be duration. In /laddɑ̃/ vs. /ladɑ̃/, for example, the critical duration is closure duration. (The duration of the vowel /a/ may also be larger before /d/ than /dd/ because of a different syllabic affiliation.) Now, if the French listeners can distinguish these pairs, can they do it based on their sensitivity to durational differences per se (Do they discriminate short vs. long voiced closure?), or to variations in the beat given by successive major acoustic/articulatory events? The latter account would engage some kind of online tracking of rhythm, such as defined by syllabic durations, vowel-to-vowel timing, or, more generally, inter-gestural timing.

In the present study, we begin with testing the first aforementioned possibility that listeners can discriminate intrinsic durations of voiced closure durations. To this end, we use Tashlhiyt Berber (henceforth, TB) minimal pairs, that are non-lexical for French listeners, with word-initial contrasts such as /b/-/bb/. The French performance is compared to that of native speakers of TB, who serve as a reference for optimal performance, in a classic cross-language design. Additionally, we test listeners on word-initial /s/-/ss/ and /f/-/ff/ contrasts, thereby manipulating the energetic content of the duration contrast. Finally, we use a word-final /t/-/tt/ contrast, which might begin to shed some light on an inter-gestural timing account of the capacity to detect gemination.

2. DISCRIMINATION EXPERIMENT

Natural utterances of TB minimal-pair words for the singleton-geminate contrast were used in a cross-language AXB discrimination test comparing native TB speakers and naïve French speakers.

2.1. Method

2.1.1. Participants

Twelve French native speakers, students or teachers at Paris 3 University (aged 21 to 57, mean 33.4, SD 13 years), and 23 TB native speakers, students at Ibnou Zohr University in Agadir (aged 19 to 37, mean 26.1, SD 4.9 years), volunteered to participate in the experiment. French and TB participants were tested in Paris and Agadir, respectively. None of the 12 French participants had any exposure to Tashlhiyt or a language using word-initial geminate-singleton contrasts. None of the French and TB participants reported hearing deficit nor any kind of language impairment.

2.1.2. Stimuli and design

Eight geminate-singleton contrasts were used: three contrasted word-initial voiced stops (*bi-bbi, diR-ddiR, and gar-garr*), another three contrasted word-final /t/ and /tt/ (*fit-fitt, hat-hatt, and jut-jutt*), and two contrasted word-initial voiceless fricatives (*fit-ffit, and siR-ssiR*). There were thus a total of 16 items. Four repetitions of each, produced by a TB native speaker, were retained as experimental stimuli. Acoustic measurements were run on the retained stimuli. As expected, the clearest cue to gemination is durational. The critical durations of the stimuli are summarized in Table 1.

Table 1: Mean durations of prevoicing, constriction, and closure durations for word-initial voiced stops, fricatives, and word-final stops, respectively (in ms).

	example	singleton	geminate
prevoicing	*bi-bbi*	70	211
fricative	*siR-ssiR*	130	252
final closure	*hat-hatt*	72	211

These duration differences all are significant at the $p<.00001$ level. They are accompanied by more subtle differences, some of which clearly reach significance. For instance, in the /t/ coda series, the longer closure for /t:/ than /t/ is partly compensated by shorter onset consonant (-22%) and initial vowel (-29%), *p*s<.001. In the voiced stop onset series, both the mean energy and F0 of the voiced murmur are higher (~+8%) in singleton than geminate consonants, *p*s<.005. Finally, there is a marginal trend for geminate fricatives to have a lower mean HNR than singleton fricatives. The same trend is found with the vowel following a geminated voiced stop (*p*s<.08).

Each contrast was presented 4 times in each of the 4 possible AXB orders so that the stimuli appeared equiprobably in all within-triplet positions. There were thus 128 trials for the eight contrasts under scrutiny. These trials were part of a larger design including 128 other trials on eight other TB contrasts, for which results are reported elsewhere. These test trials were presented in blocks of 16 trials and were preceded by 10 training trials on five different contrasts from those used in the test trials (*daR-tarr, taR-darr, kijji-gijji, tid-ttid, and jutid-juttid*).

2.1.3. Procedure

Participants were tested individually in a quiet room and received the speech stimuli through professional quality covering headphones. On each AXB trial, participants were presented with three stimuli and had to indicate whether second item X matched better the first or the third stimulus, by pressing the response key labeled '1' or '3'. The inter-stimulus (offset to onset), inter-trial, and inter-block intervals were set to 1 s, 4 s, and 9 s, respectively. Response times were measured from the onset of the X stimulus. The experiment was run using the DMDX software [5].

2.2. Results

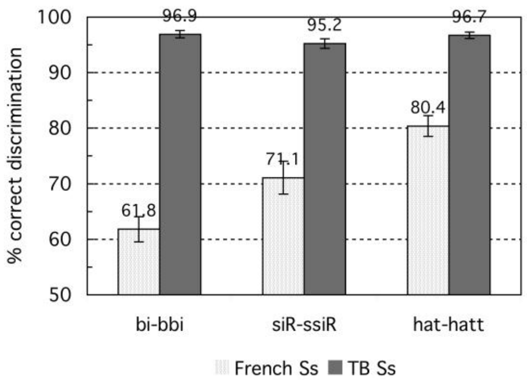
2.2.1. Correct discrimination rate

The TB participants performed near ceiling on all contrasts, as expected. The French participants performed the most poorly on the word-initial voiced stop contrasts (henceforth, D contrasts), less poorly on the word-initial voiceless fricative contrasts (henceforth, S contrasts), and better still on the word-final /t/-/tt/ contrasts (henceforth, t# contrasts). Table 2 shows the results detailed by contrast. An analysis of variance was run on these data, with Subject as a random variable, rate of correct discrimination as the dependent variable, Language as a between-subject factor (French vs. TB), AXB trial Pattern (X=A vs. B), trial Target (X=singleton vs. geminate), and Contrast type (the three types under scrutiny) as within-subject factors. The structural factor Pattern did not reach significance overall, as well as for either French or TB, and is not discussed further. The factor Target was highly significant for French, with a better performance when X in AXB was geminate than singleton (77.1 > 65.1%), $F(1,11)=36.31$, $p<.0001$. No such trend was found with TB subjects.

Table 2: Discrimination rate data detailed by contrast (standard deviations within parentheses).

	French Ss	Tashlhiyt Ss
contrast	%correct (SD)	%correct (SD)
bi-bbi	63.5 (27.8)	97.6 (8.3)
diR-ddiR	62.0 (26.3)	97.3 (9.4)
gar-ggar	59.9 (28.6)	95.9 (11.3)
fit-ffit	70.8 (30.2)	94.8 (13.1)
siR-ssiR	71.4 (27.8)	95.7 (10.2)
fit-fitt	88.0 (16.3)	95.4 (11.1)
hat-hatt	82.8 (19.4)	97.0 (8.2)
jut-jutt	70.3 (27.1)	97.8 (8.0)

Figure 1: French vs. TB correct discrimination rate data for the three types of contrasts (standard errors as positive and negative error bars).

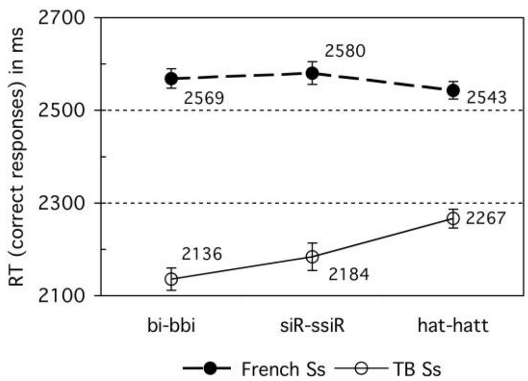


The main factor Language was highly significant, $F(1,33)= 163.12$, $p<.0001$, reflecting the better TB than French performance overall, as well as for each type of contrast (*p*s<.0001). The Contrast factor was also significant overall, $F(2,66)= 9.28$, $p<.0005$, but, as suggested by the strong Language x Contrast interaction, $F(2,66)= 18.31$, $p<.0001$, significant only for French, $F(2,22)= 10.96$, $p<.0005$, not TB, $p=.156$. Indeed, as seen in Figure 1, performance is near ceiling for TB participants for all contrasts, whereas it varies with contrast-type for French ones: ~62<71<80% for the D, S, and t# contrasts, respectively. (All the differences but the latter one are significant at the $p<.05$ level or better; 71<80 is marginally significant, $F(1,11)= 3.98$, $p=.071$; 62% was above chance level, $p<.05$.)

2.2.2. Correct discrimination rate

Figure 2 shows the RT data for correct responses. Note that RT values were measured from the onset of the second stimulus (X); RTs from the onset of the third stimulus (B) would be shorter by 1500 ms in average.)

Figure 2: French vs. TB response time data for the three types of contrasts (standard errors as positive and negative error bars).



The raw RT data was cleaned up by discarding RT values longer than 3.5 s (more than 2 s after B's onset) or shorter than 1.5 s (before B's onset). About 0.6% of the French RT data was so removed and 0.1% of the TB data. An analysis of variance was run on the cleaned-up RT data with the same factors as for the discrimination rate data. Both the Pattern and Target structural factors were significant (*p*s < .005), reflecting the same two trends in both groups of subjects: shorter RTs for "recency" (X=B) than "precedence" (X=A) trials (French subjects: 2499<2629 ms; TB subjects: 2161<2230 ms); shorter RTs for geminate than singleton target (French subjects: 2524<2604 ms; TB subjects: 2179<2212 ms). Turning now to main effects, Language was highly significant, $F(1,33)= 14.68$, $p<.0005$, with TB much faster than French listeners by about 368 ms. The TB advantage held for all three types of contrast (*p*s<.005). However, as suggested by a significant Language x Contrast interaction, $F(2,66)= 4.02$, $p<.05$, this advantage varied across contrasts: as can be seen in Fig. 1, it was smaller for the "t#" contrasts. RTs did not vary significantly across contrast types for French, $F(2,22)=1.039$, n.s., but did so for TB participants,

$F(2,44)= 5.64$, $p<.01$, with the shortest RTs for the 'D' contrast. In the correct discrimination rate and RT data, the trend for an inverse correlation between rate and RT found in other studies (e.g., [6]) is rather weak in the French data reduced along the eight contrasts used, $r(6)= -0.59$, one-tailed $p=.061$.

3. GENERAL DISCUSSION

In this study, we found that French listeners can hardly use voicing murmur duration as a cue to distinguish *bi* and *bbi*. The cross-language pattern suggests phonological deafness for this contrast, as the early work by Abramson and Lisker [2] would suggest from a phonemic repertoire perspective: French listeners have trouble discriminating within-category differences such as prevoicing differences. Likewise, French listeners encounter substantial difficulty with differences in word-initial fricative duration, in spite of a notably greater audibility for friction noise than voicing murmur. Common to these two types of acoustic contrasts, which French listeners have difficulty to perceive, is that they occur in word-initial position. French listeners comparatively have less difficulty to discriminate the duration difference in final position between *fit* and *fitt*, even though this duration is filled with silence, that is, the critical acoustic object with respect to duration is not audible. Why is that situation easier for French listeners? The intrinsic properties of the variable-duration acoustic object, be it silence, voiced murmur, or audible friction do not explain the pattern of performance observed. One would indeed expect the best performance with the most audible acoustic objects: S > D > t#. Contrary to this expectation, our data suggest the least audible interval conveys the best performance.

We propose that listeners do not measure acoustic object durations but instead measure time intervals between acoustic or articulatory events that are perceptually salient to them. Perceptual saliency may indeed vary cross-linguistically. For the issue at stake, it appears that French listeners have not acquired sufficient sensitivity to the acoustic or gestural cues that signal gemination in TB. For them, word-initial transition from silence to voiced murmur or even to friction noise is not salient enough to initiate the "beat." In word-final position, closure onset after a vowel seems salient enough to provide a clearly timed phonetic event, and the following stop release can provide the next clearly timed phonetic event. This account can explain the lesser difficulty encountered by French listeners with *fit-fitt* than *bi-bbi*. Our "beat" account is also supported, however anecdotically, by subjects' performance on *jutid-juttid* during the training phase: 100% TB and 83% French subjects discriminated *jutid-juttid*. That is, performance on *jutid-juttid* was roughly equivalent to that on *fit-fitt*. This is not surprising if we assume that listeners are sensitive to the beat between salient phonetic events: similar events follow each other in *jut(t)id* and fit(t): closure onset then closure release. French listeners are no more sensitive to inter-event beat when filled with very audible material such as a vowel, as suggested by Dupoux and colleagues [4], who found that French listeners can hardly discriminate *ebuzo-ebuuzo*.

To sum up, the better performance of TB than French listeners is presumably due to language-specific attunement leading to enhanced sensitivity to the acoustic or gestural cues that are relevant to gemination. The pattern of French performance on *bi-bbi*, *fit-ffit*, and *fit-fitt* suggests that listeners who did not acquire gemination can detect it only from the beat provided by salient phonetic events, be they acoustic or articulatory in nature.

4. REFERENCES

[1] Abramson, A. 1986. The perception of word-initial consonant length: Pattani Malay. *J. Int. Phon. Assoc.* 16, 8-16.

[2] Abramson, A., Lisker, L. 1970. Discriminability along the voicing continuum: Cross-language tests. *Proceedings of the 6th International Congress of Phonetic Sciences* Prague, 569-573.

[3] Abramson, A., Lisker, L. 1973. Voice-timing perception in Spanish word-initial stops. *J. Phon.* 1, 1-18.

[4] Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., Mehler, J. 1999. Epenthetic vowels in Japanese: A perceptual illusion? *J. Expt. Psych.: Human P. and P.* 25, 1568-1578.

[5] Forster, K., Forster, J. 2003. DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers* 35, 116-124.

[6] Hallé, P., Best, C. 2007. Dental-to-velar perceptual assimilation: A cross-linguistic study of the perception of dental stop+/l/ clusters. *J. Acoust. Soc. Am.* 121, 2899-2914.

[7] Meisenburg, T. 2006. Fake geminates in French: a production and perception study. *Conference of Speech Prosody* Dresden, Germany.

[8] Ridouane, R. 2007. Gemination in Tashlhiyt Berber: An acoustic and articulatory study. *J. Intl Phon. Assoc.* 37, 119-142.

[9] Serniclaes, W. 1987. *Etude Expérimentale de la Perception du trait de Voisement des Occlusives du Français*. Ph.D. dissertation, Université Libre de Bruxelles.