In: Bell-Berti, F., & Raphael, L. J. (1995). Producing Speech: Contemporary Issues. For Katherine Safford Harris. AIP Press: New York.

8 Intermediate Values of Voice Onset Time

Lawrence J. Raphael
Herbert H. Lehman College, CUNY, The Graduate School, CUNY, and Haskins Laboratories

Yishai Tobin
Ben-Gurion University of the Negev and University of Tel Aviv

Alice Faber
Haskins Laboratories and Wesleyan University

Tova Most
University of Tel-Aviv

H. Betty Kollia
Brooklyn College, CUNY, and Haskins Laboratories

Doron Milstein
The Graduate School, CUNY

This potpourri of authors includes not only Katherine Harris' colleagues and her students, but also her students' students. We are all grateful to her for teaching us to think clearly, to explain data rather than explaining them away, and to adopt a broad perspective even when conducting a narrow experiment.

Beginning with the publication of Lisker and Abramson's (1964) landmark paper, it has become conventional to assign the Voice Onset Time (VOT) of a

prevocalic stop consonant to one of three categories commonly found in the natural languages of the world: (1) voicing lead (approximately –30 ms. or more VOT); (2) zero onset/short-lag (approximately 0 to +30 ms VOT); and (3) long-lag (approximately +50 ms or more VOT). In general, these categories of VOT correspond, respectively, to the phonetic descriptions of stop consonants as (1) voiced, unaspirated, (2) voiceless, unaspirated, and (3) voiceless, aspirated.[1]

Lisker and Abramson's data also revealed that in languages employing a single phonological voicing opposition, speakers realize stops phonetically using VOTs from adjacent categories (i.e., from the voicing lead and zero onset/short-lag categories, or from the zero onset/short-lag and long-lag categories) and not from the extreme categories (i.e., voicing lead and long-lag). The generality of this observation, however, may require some modification, as suggested by some recent studies (Caramazza, Yeni-Komshian, Zurif, & Carbone, 1973; Keating, Mikos, & Ganong, 1981; Recasens, 1985; Yeni-Komshian, Caramazza, & Preston, 1977), including the one we present here.

Our initial studies of VOT in Hebrew, more than a decade ago (Raphael & Tobin, 1983; Raphael, Tobin, Faber, & Most, 1989; Raphael, Tobin, & Most, 1983), were, in fact, prompted by the conflicting opinions and data provided by phoneticians and phonologists who have studied Modern Hebrew. Rosén's (1966) phonological study, for example, maintains that Hebrew /ptk/ are never aspirated, even in syllable-initial position, which is the expectation for stops in languages like Spanish and Italian, which draw these stops from the zero onset/short-lag category of VOT. Chayen (1973) is somewhat more equivocal, appealing to the so-called fortis/lenis distinction as the basis for distinguishing Hebrew /ptk/ from /bdg/. He does, however, indicate that both voicing and aspiration play a minor role in the distinction, and mentions that heavy aspiration can sometimes occur in association with syllables bearing emphatic stress. He also observes that aspiration is stereotypically associated with British- or American-accented Hebrew.

In contrast, other commentators have suggested that it is not uncommon for some degree of aspiration to be associated with Hebrew /ptk/. Blanc (1956, 1964) presents relatively narrow phonetic transcriptions of Hebrew and comments that aspiration (especially of /t/) is evidenced by a number of speakers.[2] Laufer (1972), in his synthesis of Hebrew stops, used VOT values of +20 ms for /p/ and /t/, and +30 ms for /k/. These values may be taken to represent a slight degree of aspiration. Indeed, Laufer, a native speaker of

[1]The phonetic category of voiced, aspirated stops is not distinguished by a unique range of VOT values (Lisker & Abramson, 1964).

[2]Blanc's studies established the distinction between the dialect of Modern Hebrew spoken by those with a central or eastern European linguistic heritage (Ashkenazic) and the dialect spoken by those with a Mediterranean/Middle-Eastern linguistic heritage (Sephardic). He reports that the two dialects differ in their use of aspiration, but does not amplify further.

Hebrew, reports that aspiration "may or may not follow the release of /ptk/" in his own speech.[3]

More recent, instrumental, studies of VOT in Hebrew reinforce the observations of Blanc and Laufer. Devens' (1978) study reported values of about +30 ms VOT for Hebrew /pt/ and a value of about +65 ms for /k/. Obler's (1982) study reported mean VOT values of +25.6 ms, +33.9 ms, and +63.7 ms for /ptk/, respectively, for five monolingual speakers of Hebrew. All other things being equal, these data might lead us to infer a moderate degree of aspiration accompanying the release of /p/ and /t/, and a degree of aspiration approaching that found in long-lag languages (e.g., English) for /k/. Obler's data for 10 balanced bilingual speakers of Hebrew and English present a more extreme picture, with mean VOT values of +58.0 ms, +73.8 ms, and +93.2 ms for Hebrew /ptk/, respectively. These values fall well into the ranges evidenced by native English speakers, and suggest that the stops are aspirated (at least by some of the speakers) in much the same way they might be in English.[4] Obler notes that the relatively long VOTs for Hebrew /ptk/ are accompanied by typical voicing lead values for Hebrew /bdg/, a finding reported by other investigators for bilingual speakers of voicing lead/short-lag languages who have acquired and short-lag/long-lag language as their L2 (Flege, 1987, 1991). It would appear, then, that some bilinguals, when using L2, are not constrained (or able) to draw their stops from adjacent VOT categories, using voicing-lead values from their L1 for /bdg/, and voicing-lag values either intermediate to their L1 and L2 for /ptk/ or equivalent to the values of native speakers of their L2 (in the case of balanced bilinguals).[5]

It was our hope that using a larger pool of subjects than those of the previous studies might help to resolve some of the uncertainty regarding VOT and aspiration in Hebrew /ptk/.

STUDY I

Procedure

Our subjects were 23 students at the University of Tel Aviv and Ben-Gurion University of the Negev, and seven Israelis who have lived in the New York

[3]The impression of several of the authors of this paper who are either native speakers of Hebrew or who have heard Hebrew spoken in Israel (and in the United States) is that /ptk/ are produced with a moderate degree of aspiration by many speakers, and, by some speakers, with a degree of aspiration typical of long-lag languages like English.

[4]Indeed, the mean value for the velar stop (+93.2 ms) exceeds that found by Obler for five monolingual speakers of English (89.4 ms). Moreover, the bilinguals' VOT values for /ptk/ equal or exceed those values reported by Lisker and Abramson (1964) for English.

[5]We might assume that the bilinguals maintain voicing lead in /bdg/ because native English-speaking listeners do not distinguish voicing lead from zero-onset/short-lag.

City Metropolitan area for between two and five years. All were native speakers of Hebrew. Of the 30 subjects, 10 were male and 20 were female. All spoke English fluently, but with an easily discernible accent.[6] They recorded 30 minimal pairs of Hebrew words, ten pairs for each of the three places of articulation. The minimal pair list used in Israel was arranged to provide maximal contrast, with each token of a /ptk/ word immediately preceding a token of its respective voiced cognate; the list used for recording in the United States was randomized.

The speakers' productions were analyzed on a Kay Elemetrics Sonograph, model 6061B, using a wideband filter. In addition to the standard wideband display, we generated a waveform at the top of each spectrogram. VOT's were first measured from the wideband display, using the stop burst and the first discernible vocal pulse at the level of the baseline as points of reference for /ptk/. In instances of voicing lead, VOT was measured from the burst only as far back as there was evidence of continuous vocal pulsing. Measurements were estimated to the nearest 5 ms. A second set of measurements was made from the accompanying waveform, using the burst and the onset of periodicity as points of reference. Both sets of measurements were made by each of two of the authors. Agreement between measurements was better than 90 percent. Differences were resolved by a third set of measurements performed jointly by the two authors who made the original measurements.

Results

The distribution of VOT values for the Hebrew stops recorded in Israel is described below. The mean values for these data are shown in Table 1. The VOT ranges and means for Hebrew are typical of those found in languages that use voicing lead for one phonological category of stops. These data are in general agreement with Obler's for the same sounds. The mean VOT for /b/, shown in Table 1, is –91.9 ms, with a clear mode in the data at –90 ms. Values ranged from –145 ms to –30 ms. The –90 ms mean for /d/ was accompanied by a nominal mode at –85 ms and a range of values from –135 ms to –40 ms.[7] The mean VOT for /g/ was –81.3 ms, with a nominal mode at –80 ms and a range of values from –145 ms to –25 ms.

[6]Languages other than English that were spoken by our subjects included German, Hungarian, Yiddish, Russian, and Polish.

[7]We have excluded from these data three intended productions of /d/, which we perceived as /t/: one at +5 ms and two at +10 ms. These values overlapped those for /t/—the only instance of overlap in our data.

TABLE 1. *Mean VOT values (Hebrew-speaking adults) from data recorded in Israel.*

	HEBREW (ISRAEL) MEAN VOICE ONSET TIME (ms)		
/b/	−91.9	+28.5	/p/
/d/	−90.9	+35.2	/t/
/g/	−81.3	+54.2	/k/

The data for /ptk/ are generally intermediate to those found in languages that contrast short-lag and long-lag stops. The mean VOT for /p/ is 28.5 ms, with a mode at +15 ms and a range of −5 ms to +85 ms. For /t/, the mean and mode were +35.2 ms and +30 ms, respectively, with a range of 0 to 65 ms The velar, /k/, displayed a mean VOT of +54.2 ms with a mode at +50 ms, and a range of −10 ms to +105 ms. A comparison of these data with those of Obler's study reveals general agreement between our bilingual subjects and Obler's monolinguals. In contrast, Obler's (balanced) bilingual subjects, as we have mentioned, displayed VOT values much closer to those of native English speakers and, thus, considerably greater than those of our (not-balanced) bilinguals. Devens' data, also for bilinguals,[8] is in close agreement with ours (and with Obler's for monolingual speakers).

The distributions of VOT values for the 7 US-based subjects were similar in every regard to those of the Israeli-based subjects. The mean VOT values for these subjects are shown in Table 2. We have pooled the values for the two sets of subjects in Table 3.

TABLE 2. *Mean VOT values (Hebrew-speaking adults) from data recorded in the United States.*

	HEBREW (USA) MEAN VOICE ONSET TIME (ms)		
/b/	−82.2	+25.4	/p/
/d/	−91.5	+34.2	/t/
/g/	−77.4	+66.1	/k/

[8]Devens (1980) does not specify whether or not her 10 subjects were balanced bilinguals. The first language of some of her subjects might have been Arabic, and they may also have been speakers of English or French, as well as Hebrew.

TABLE 3. *Mean VOT values (Hebrew-speaking adults) pooled across data recorded in Israel and in the United States.*

	HEBREW (ALL SS) MEAN VOICE ONSET TIME (ms)		
/b/	−89.8	+28.3	/p/
/d/	−90.9	+35.6	/t/
/g/	−79.2	+55.5	/k/

The values in Tables 1–3 indicate that bilingual speakers of Hebrew produce /bdg/ with VOT values that are quite typical of languages that employ voicing lead in a two-way stop opposition. The values for /ptk/, however, do not provide a good fit with the voiceless inaspirates of such languages, nor with the voiceless aspirates of languages like English, which distinguish short-lag from long-lag stops. That is, they are somewhat closer to the mean values of the former group of languages, but appear to constitute an independent set of means and ranges.

One possible explanation for the distribution of VOT values that we found is that modern Hebrew is in the midst of a shift from short-lag to long-lag values for /ptk/. Another possibility is that VOT values in modern Hebrew were much like they are now when the language was revived about a century ago. One possible way of gaining insights about these alternatives might be to investigate the VOT values of children born in Israel who speak Hebrew as a first language. It is possible that, if a shift in values were in progress, the children might evidence greater values of voicing lag than adults for the same stop consonants. Accordingly, we turned our attention to a younger generation of speakers.

STUDY II

Procedure

Our subjects were 21 10- to 11-year old children (range 10.0-11.1 years) who were students at Be'eri elementary school in Be'er Sheva, Israel. Eight of the students were male and 13 were female. All were native speakers of modern Hebrew. Four of the students had some knowledge of another language (Romanian—two children; French and English—one each), although none could be classified as skilled bilinguals. The children recorded a list of 30 monosyllablic minimal pairs containing some of the words read by the adults in Study I. Their first reading was from a randomized list, the second from a list in which each minimal pair constituted a sequence, the words beginning with /ptk/ preceding those beginning with /bdg/. The methods of analysis and measurement were the same as those used in Study I, with the exception that measurements were made on a Kay Elemetrics Digital Sonagraph Model 7800.

Results

Because of the well-known variability in speech production evidenced by children, relative to adults, we expected that the data for the children would be somewhat less systematic than those from Study I. Indeed, the ranges of the children's VOT values for each of the stop sounds were greater than those of the adults, and, as a result, there was more overlap between the VOT values for each pair of cognates.[9] There were no clear modes in the distribution of VOT values for /bdg/ and the modes for /ptk/ were consistently smaller than the means for those sounds. Table 4 shows the mean VOT values for the stops read in words from the randomized list, the minimal pairs read as pairs, and the two conditions pooled.

As shown in Table 5, the children produced /bdg/ with values of voicing lead that were somewhat lower than those of the adults in all three places of articulation. Nonetheless, the mean voicing lead values are quite typical for languages which employ this VOT category. The means of the children and the adults for /ptk/ are extremely similar for the labial and velar stops, showing differences of 0.3 ms and 1.4 ms, respectively. The means for /t/ differ by about 7 ms. The children's VOT means for /ptk/, like those of the adults, fall at values intermediate to the short- and long-lag values typical of many languages.

TABLE 4. *Mean VOT values (Hebrew-speaking children) from randomized word lists, minimal pair lists, and data pooled across both conditions.*

		MEAN VOICE ONSET TIME (ms)			
	RANDOM	−78.4	+25.1	RANDOM	
/b/	PAIRS	−75.2	+28.2	PAIRS	/p/
	POOLED	−76.9	+26.9	POOLED	
	RANDOM	−67.7	+22.5	RANDOM	
/d/	PAIRS	−76.9	+28.3	PAIRS	/t/
	POOLED	−72.5	+25.1	POOLED	
	RANDOM	−80.7	+65.4	RANDOM	
/g/	PAIRS	−70.9	+55.6	PAIRS	/k/
	POOLED	−75.4	+60.6	POOLED	

[9]The number of cases of overlap was still quite small in relation to the total number of measurements: 75 out of more than 2400 measurements.

TABLE 5. *Comparison of mean VOT values for minimal pair word lists produced by Hebrew-speaking adults and children.*

	MEAN VOT (PAIRED WORDS)					
	/b/	/p/	/d/	/t/	/g/	/k/
ADULTS	–91.9	+28.5	–90.9	+35.2	–81.3	+54.2
CHILDREN	–75.2	+28.2	–76.9	+28.3	–70.9	+55.6

DISCUSSION

The existence of VOT values intermediate to the short- and long-lag categories has, by this time, a considerable amount of precedent. Such an intermediate category has long been known to exist in Korean (Lisker & Abramson, 1964), although it exists between the short-lag and long-lag stops of that language. Caramazza et al. (1973) studied VOT in French bilinguals in Canada. The bilingual, native French speakers in their study produced /ptk/ in French with intermediate VOT values.[10] Our own study of native speakers of Puerto Rican Spanish (Raphael et al., 1983) and Greek (Kollia, 1993) revealed mean VOTs for /ptk/ that were slightly smaller than those reported here for speakers of Hebrew, but which fall into the intermediate category. Similar results have also been found for Polish (Keating et al., 1981), Catalan (Recasens, 1985), and Lebanese Arabic (Yeni-Komshian et al., 1977). The mean VOTs from these studies for /ptk/, and from the present study, are presented in Table 6.

TABLE 6. *Mean VOT values for /p t k/ from studies of seven languages.*

STUDY	LANGUAGE	/p/	/t/	/k/
Present Study (adults)	Hebrew	28.3	35.6	55.5
Obler, 1982 (Monolinguals)	Hebrew	25.6	33.9	63.7
Raphael et al., 1983	Spanish	20.2	27.6	39.3
Caramazza et al., 1973	Québec French	20	28	35
Keating et al., 1981	Polish	21	28	52
Recasens, 1985	Catalan	23	27	47
Yeni-Komshian et al., 1977	Lebanese Arabic	——	25	30
Kollia, 1993	Greek	18.5	27.1	49.2

[10]They also found a considerable amount of overlap between the two classes of French stops for both bilingual and monolingual speakers of French. This overlap, coupled with their perceptual data for French-speaking monolinguals led them to conclude that VOT is not an "important variable for voicing distinctions...in Canadian French." We make no such claim for Modern Hebrew.

It is tempting to attribute the existence of an intermediate category of VOT to effects of bilingualism. (See Flege, 1987, 1991 for some compelling reasons why this can be so.) In most of the studies we have cited, including ours, many or all of the speakers were bilingual. For instance, Kollia's (1993) subjects were Greek students tested in New York; all spoke English, but Greek was clearly dominant for them. Moreover, studies of English spoken by non-native speakers have presented data with VOT values that can be termed intermediate (Flege & Eefting, 1987; Gass, 1984). However, it seems clear that bilinguals are not the only speakers who produce intermediate VOT values. Monolingual speakers of Hebrew (including the children in the present study) and Obler's (1982) adult subjects, Polish (Keating et al., 1981) and Catalan (Recasens, 1985) have produced such values. The subjects in the study of Lebanese Arabic (Yeni-Komshian et al., 1977) were either students who were studying English or employees of the American University of Beirut, whose "knowledge of English was either very limited or nonexistent." Moreover, even though many of our Hebrew-speaking subjects were bilingual, the other language(s) they spoke, whether natively or as an L2, were not in all cases languages which employ long-lag stops. We do not mean to imply that the linguistic pressures exerted on speakers who are bilingual cannot or do not directly influence VOT or any other aspect of language. But such pressures cannot serve as a general explanation for all of the VOT data we have presented here.

Finally, considering the ranges and means for VOT that we and others have reported, as well as the fact that both child and adult monolinguals evidence intermediate values of VOT, it seems to us that any attempts to link VOT categories causally to auditory sensitivities (e.g., Diehl & Kluender, 1989; Pisoni, 1977) are largely unwarranted. Rather, it appears more likely that the perceptual mechanisms, especially of children, are tuned to the productions of skilled speakers, no matter what category of VOT those skilled speakers may employ.

ACKNOWLEDGMENTS

The research reported herein and the preparation of the manuscript were supported by grants HD-01994, DC-00403, and DC-00121 from the National Institutes of Health to Haskins Laboratories. We are grateful to Loraine Obler for discussion of the issues treated in this paper and for her helpful comments on an earlier version thereof.

REFERENCES

Blanc, H. (1956). A selection of Israeli Hebrew Speech. *Leshonenu, 21*, 33–39.

Blanc, H. (1964). Israeli Hebrew texts. *Studies in Egyptology and linguistics in honor of H. J. Polotsky* (pp. 132–152). Jerusalem: Israel Exploratory Society.

Caramazza, A., Yeni-Komshian, G., Zurif, E., & Carbone, E. (1973). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *Journal of the Acoustical Society of America, 54*, 421–428.

Chayen, M. (1973). *The phonetics of Modern Hebrew*. The Hague: Mouton.

Devens, M. (1978). *The phonetics of Israeli Hebrew: Oriental vs. General Israeli Hebrew*. Unpublished doctoral dissertation, UCLA.

Devens, M. (1980). Oriental Israeli Hebrew: A study in phonetics. *Afroasiatic Linguistics, 7*, 127-142.

Diehl, R. L., & Kluender, K. R. (1989). On the objects of speech perception. *Ecological Psychology, 1*, 121–144.

Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics, 15*, 47–65.

Flege, J. E. (1991). Age of learning affects the authenticity of voice onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America, 89*, 395–411.

Flege, J. E., & Eefting, W. (1987). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics, 15*, 67–83.

Gass, S. (1984). Development of speech perception and speech production abilities in adult second language learners. *Applied Psycholinguistics, 5*, 51–74.

Keating, P. A., Mikos, M. J., & Ganong, W. F. III. (1981). A cross-language study of range of voice onset time in the perception of initial stop voicing. *Journal of the Acoustical Society of America, 70*, 1261–1271.

Kollia, H. B. (1993). Segmental duration changes due to variations in stress, vowel, place of articulation, and voicing of stop consonants in Greek. *Journal of the Acoustical Society of America, 93*, 2298(A).

Laufer, A. (1972). *Synthesis by rule of a Hebrew dialect*. Unpublished doctoral dissertation, University of London.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: acoustical measurements. *Word, 20*, 384–422.

Obler, L. (1982). The parsimonious bilingual. In L. Obler & L. Menn (Eds.), *Exceptional language and linguistics* (pp. 339–346). New York: Academic Press.

Pisoni, D. B. (1977). Identification and discrimination of the relative onset of two component tones. *Journal of the Acoustical Society of America, 61*, 1352–1361.

Raphael, L. J., & Tobin, Y. (1983). Perceptual and acoustic studies of voice onset time in Hebrew. In A. Cohen & M. P. R. van den Broeke (Eds.), *Abstracts of the Tenth International Congress of Phonetic Sciences* (p. 516). Dordrecht, The Netherlands: Foris Publications.

Raphael, L. J., Tobin, Y., Faber, A., & Most, T. (1989). Voice onset time in Hebrew: acoustic and perceptual studies. Paper presented at the International Conference on Linguistic Approaches to Phonetics, Ann Arbor Michigan.

Raphael, L. J., Tobin, Y., & Most, T. (1983). Atypical VOT categories in Hebrew and Spanish. *Journal of the Acoustical Society of America, 74*, S89A.

Recasens, D. (1985). *Estudis de fonètica experimental del Català Oriental Central.* Barcelona: Publicacions de l'Abadia de Monserrat.

Rosén, H. B. (1966). *A textbook of Israeli Hebrew*. Second, corrected edition. Chicago: University of Chicago Press.

Yeni-Komshian, G., Caramazza, A., & Preston, M. S. (1977). A study of voicing in Lebanese Arabic. *Journal of Phonetics, 5*, 35–48.