OFFPRINT FROM

HEBREW STUDIES

A Journal Devoted to
Hebrew Language and Literature

Volume XXXIX, 1998

Laufer (1998), "Voicing in Contemporary Hebrew..."

Published at the University of Wisconsin–Madison by
The National Association of Professors of Hebrew
in American Institutions of Higher Learning

איגוד הפרופסורים לעברית בארצות־הברית

VOICING IN CONTEMPORARY HEBREW IN COMPARISON WITH OTHER LANGUAGES[1]

Asher Laufer
The Hebrew University

As some scholars believe that voicing is only an accompanying, non-phonemic feature and that the essential characteristic in Hebrew is the force of articulation, we were encouraged to check the phonetic voicing in Hebrew. Since final sounds cannot be measured with an ordinary VOT method, we used also the "Voice into Closure" (ViC) method.

Measurements of fourteen Hebrew subjects revealed that in every voiced stop the phonetic voicing extends longer than in its voiceless counterpart. Comparing our VOT findings to those in twelve other languages, revealed that in a "voiced" sound there is always longer voicing than in its voiceless counterpart. Some characteristics that differentiate between languages are reported (e.g., Hebrew speakers resemble Spanish and Polish speakers with a large span of VOT, in comparison to the nine other languages checked).

Perception tests of synthesized words (on English, Spanish, Thai, and Hebrew listeners) demonstrated that the voice timing cue by itself suffices to differentiate between the voiced-voiceless categories. While the hypothesis of "force" could not be supported, "voice timing" was found to be an actual physical feature, and it is reasonable to assume that this feature is the main cause of a categorical differentiation in all the languages.

INTRODUCTION

In the fifteenth century, Leonardo da-Vinci is said to have demonstrated voicing using the lungs, trachea, and larynx of a human cadaver,[2] but the role of the vocal folds in speech was discovered by the Frenchman, Antoine Ferrein. In 1741, with a larynx of a dog, and afterwards with a human cadaver, he demonstrated the vibration of the vocal folds to the French Science Academy, explaining that these vibrations are essential for speech.

---

[1] An early version of this paper was published in Hebrew in A. Laufer, "Voicing in Colloquial Hebrew" in *Leshonenu* 57 (1993) 299–342. However, this paper uses a different method for measuring final stops, adds some more data and has more extensive discussions. I wish to thank the Institute of Jewish Studies and the Federman Foundation for their support in this research. For this research I also made use of four seminar papers written under my tutelage during 1991–1992, and I would like to thank the four students who conducted them: Ms. Na'ama Bahat, Ms. Vardit Goldmacher, Ms. Vered Levkovits, and Ms. Ronit Tsachar. I would also like to thank Mr. Ilan Makabrit for assisting me with the computer programming and the drawing of the figures and Mr. Doron Modan for his help in editing. Thanks are due also to my anonymous reviewers for their perceptive comments leading to clarification of several points.

[2] H. M. Kaplan, *Anatomy and Physiology of Speech* (Second edition; New York: McGraw-Hill, 1971), pp. 251–253.

Since Ferrein's time most linguists believed that the feature differentiating pairs like the English "*pill/bill*" was the physiological vibration of the vocal folds accompanying the articulation of the voiced sound /b/ and the lack of such vibrations during the voiceless sound /p/. However, in our century, linguists started to doubt whether phonetic voicing had a phonemic role in the languages of the world.

The search for a criterion to substitute for "voicing" was probably caused indirectly by Henry Sweet's observation that in certain contexts in English, in stops and in fricatives that were considered to be voiced, the vibration of the vocal folds was minimal or even completely absent.[3] He pointed out that in an initial stop like /g/ in English there is hardly any voicing.[4] His subjective hearing was supported experimentally by researchers who found that in English initial and final stops there may indeed be little or even no vibration of the vocal folds, although they are considered voiced stops.[5]

Such observations could have caused a dilemma, since experiments proved that a difference exists in pairs like /b-p, d-t, v-f, z-s/, even in initial or in final positions. Thus, for example, English listeners can differentiate between words like:

few/view	*pill/bill*	*ton/done*	*cane/gain*	*chin/gin*
proof/prove	*cup/cub*	*built/build*	*back/bag*	*ice/eyes.*

If a phonological difference exists between members of such pairs and if that difference is not the physiological "voicing," what causes it?

Scince the mid-part of the century, linguists started to use new criteria for describing the difference between pairs like those above, using vague criteria such as fortis/lenis, strong/weak, intensive/non-intensive, or tense/lax.[6] Although there is no consensus regarding the terminology, we shall refer henceforth to all these as fortis/lenis and call the new criterion "force."

---

[3] H. Sweet, *A Handbook of Phonetics, Including a Popular Exposition of the Principles of Spelling Reform* (Oxford: Clarendon Press, 1877), pp. 75–80.

[4] Henry Sweet, *A Primer of Spoken English* (Oxford: Clarendon Press, 1932), pp. 8–9.

[5] E.g., P. Roach, *English Phonetics and Phonology: A Practical Course* (Cambridge: University Press, 1983), p. 31; P. Roach, *English Phonetics and Phonology: Tutor's Book* (Cambridge: University Press, 1983), p. 22; L. B. Crane, E. Yeager, and R. L. Whitman, *An Introduction to Linguistics* (Boston: Little Brown, 1981).

[6] E.g., R. Jakobson, G. Fant, and M. Halle, *Preliminaries to Speech Analysis* (MIT, Acoustic Laboratory, Cambridge, Mass., Technical Report No 13, 1952; also MIT Press, 1963), pp. 26–40; N. Chomsky and M. Halle, *The Sound Pattern of English* (New York: Harper and Row, 1968), pp. 324–328.

A survey of the linguistic literature of our century reveals that the terms fortis/lenis have been used (a) as labels for the phonological contrast between cognates such as /p–b/, /s–z/, etc., and (b) as descriptive terms for the phonetic basis of that contrast. There is a significant problem with (b), arising from the fact that the empirical evidence in this area is not supportive. In our survey, "'*voiced / voiceless*' *against* '*lenis / fortis*,'" we bring forth the hypotheses of various scholars and analyze the physiological, aerodynamic, or acoustic components they stated to comprise the criterion of "force."[7]

We could not find any convincing phonetic evidence to show that the criterion of force differentiates pairs of sounds like [b–p, d–t, g–k, v–f, z–s]. Experiments in the field that were carried out in various languages have not provided a phonetic basis for hypotheses that "fortis" sounds are **more accurate** or **more tense** or that they are produced through **greater muscle tension** or **greater lung pressure** than their counterparts thought to be "lenis." On the contrary, there is experimental evidence that denies it.

It has also been claimed that **aspiration, supraglottal air pressure**, and **duration** are tied to the feature "fortis," but such a tie is artificial. Aspiration is not a separate factor from voicing, but a direct consequence of the laryngeal timing: Aspiration is the result of the delay of the vocal fold vibration after the release of the sound and the escape of air through the glottis.[8] The "voicing" or the position of the glottis during the production of voiced and voiceless sounds also explains the difference in

---

[7] A. Laufer, "'voiced / voiceless' against 'lenis / fortis,'" in *Hebrew Linguistics* 39 (1995) 41–65 [Heb., English abstract: pp. VI–VII]. We can mention here only a few of the studies discussed in the survey: Jakobson *et al.*, *Preliminaries to Speech Analysis*; E. Fischer-Jørgensen and T. Hansen, "An Electrical Manometer and its Use in Phonetic Research," *Phonetica* 4 (1959) 43–53; L. Lisker and A. S. Abramson, "A Cross-language Study of Voicing in Initial Stops: Acoustical Measurements," *Word* 20 (1964) 384–422 [reprinted in *Readings in Clinical Spectrography of Speech*, eds. J. Baken and R. G. Daniloff (1991), pp. 247–285, Singular and Kay]; L. Lisker and A. S. Abramson, "Distinctive Features and Laryngeal Control" *Language* 47 (1971) 767–785; K. S. Harris, T. Gay, G. F. Lysaught, and M. M. Schvey, "Some Aspects of the Production of the Oral and Nasal Labial Stops," *Language & Speech* 8 (1965) 135–147; A. Malecot, "Mechanical Pressure as an Index 'Force of Articulation'" *Phonetica* 14 (1966) 169–180; N. Chomsky and M. Halle, *Sound Pattern*; G. Fant, *Speech Sounds and Features* (Cambridge: M.I.T. Press, 1973); L. J. Raphael and F. Bell-Berti, "Tongue Musculature and the Feature of Tension in English Vowels" *Phonetica* 32 (1975) 61–73; J. C. Catford, *Fundamental Problems in Phonetics* (Bloomington: Indiana University Press, 1977); R. Daniloff, G. Schuckers, and L. Feth, *The Physiology of Speech and Hearing* (Englewood Cliffs, NJ: Prentice-Hall, 1980); J. Laver, *The Phonetic Description of Voice Quality* (Cambridge: Cambridge University Press, 1980) 95–99; I. R. Titze and R. C. Schefer, eds., *Vocal Fold Physiology: Biomechanics, Acoustics & Phonatory Control* (Denver, CO, 1983). For more details see the rich bibliography in the survey itself.

[8] E.g., J. C. Catford, *Fundamental Problems*, p. 115; A. S. Abramson, "Laryngeal Timing in Consonant Distinction" *Phonetica* 34 (1977) 295–303.

the **supraglottal air pressure** in a natural and simple fashion. The essence of voicing does not permit a rise in the supraglottal air pressure. In the same way, **short duration** of a sound that is considered voiced cannot be linked with the criterion of "fortis," but it can at least partially be linked to the activity of the vocal folds. Supraglottal pressure must be lower than sub-glottal pressure in order to enable voicing, and one of the ways to achieve this is by shortening the duration of the voiced sound.

Some of the phonetic features which have been claimed to be tied to the criterion of force do not actually exist, while features that do exist do not have an obvious connection to "force." Instead, there is a causal connection between the features that do exist and physiological voicing. "Force" is not, therefore, a phonetic criterion, while phonemic voicing should be based on phonetic "voice timing."

Phonetic "voicing" is the vibration of the vocal folds, while the phonological criterion of "voicing" is determined by "relative phonetic voicing." When we say that a sound is "voiced," we do not, therefore, necessarily mean that phonetic voicing is actually occurring during the whole production of the sound but that there is relatively more phonetic voice during the production of the "voiced" sound in comparison to its "voiceless" counterpart. Such an explanation of "voice timing" solves the dilemma raised by Sweet. We shall return to "voice timing" in detail after giving the results of the measurements of Hebrew voicing.

Aims of This Research

As some scholars believed that the essential characteristic in Hebrew is the force of articulation and that voicing is only an accompanying non-phonemic feature,[9] we wanted to check if voicing is a reliable and sufficient cue to distinguish sounds in contemporary Hebrew. This was done in three stages: In the first stage, we confirmed experimentally that Hebrew really has a phonological feature we could name "voicing." In the second stage, we examined whether that phonological feature can be connected to the acoustic feature of voice timing. In the third stage, we set out to find whether the acoustic attributes found in the second stage sufficed to account for Hebrew listeners' perception of "voicing."

---

[9] D. Tene, "The Measured Duration of Hebrew Vowels (kymographical study)" *Leshonenu* 26 (1962) §§ 4.6, 5.3 220–268 (Heb.); M. J. Chayen, "Pronunciation of Israeli Hebrew" *Leshonenu* 36 (1972) 288, 290 (Heb.); M. J. Chayen and Z. Dror, *Introduction to Hebrew Transformational Grammar* (Heb.; Tel Aviv: University Publications, 1976), p. 294.

Research Limits

On the one hand, the feature of phonetic voicing is universal and has been observed in all languages examined.[10] In most sounds, vocal fold vibrations cause the acoustic input and can therefore be considered the main cause for our ability to hear the sounds.[11] Vocal fold vibrations are the only acoustic input in vowels, in glides, and in most nasals and liquids. They are one of two sources in nonsonorant voiced sounds. On the other hand, few researchers have found experimentally that vocal fold vibrations do not always accompany sounds that are perceived as "voiced."[12] This phenomenon was discovered only in a few languages, only in a few sounds (mainly stops), and only in certain environments (sometimes in initial positions, and rarely in final positions). Therefore, as the problem of voicing arose mainly with stops in initial and final positions, we limited our research to these cases.

1. VOICING IN CONTEMPORARY HEBREW

Hebrew researchers unanimously believe that Hebrew speakers have the ability to differentiate and discriminate between pairs like /p b/, /t d/ and /k g/.[13] Still, an initial perceptual experiment to test this belief should be a sound basis for the more analytical experiments to follow. Therefore, we recorded native Hebrew speakers pronouncing Hebrew utterances that contained these phonemes and could be used as minimal pairs. Contemporary Hebrew contrasts /b d g/ with /p t k/ in initial, medial and final positions, but since the problem of voicing arose only in initial and final stops,[14] we limited our perception tests to these instances. We shall elaborate on this stage, because we also use the same recordings for our next stage.

---

[10] J. E. Clark and C. Yallop, *An Introduction to Phonetics & Phonology* (Second edition; Oxford: Blackwell, 1995), p. 123; P. Ladefoged and I. Maddieson, *The Sound of the World's Languages* (Oxford: Blackwell, 1996).

[11] E.g., P. Ladefoged, *Elements of Acoustic Phonetics* (Chicago: University of Chicago Press, 1962); J. Laver, *The Phonetic Description of Voice Quality* (Cambridge: Cambridge University Press, 1980).

[12] Laufer, "'voiced / voiceless.'"

[13] E.g., see H. B. Rosen, *A Textbook of Israeli Hebrew* (Second edition; Chicago: University of Chicago Press, 1969); M. J. Chayen, *The Phonetics of Modern Hebrew* (Janua Linguarum 162; Hague: Mouton, 1973).

[14] A. Laufer, "'voiced / voiceless.'"

1.1 First Perception Test of Natural Hebrew Speech

1.1.1 Preparation of the Material for Recording and the Perception Test

Hebrew words starting and ending with /p b t d k g/ were chosen in a way that they could be used as minimal pairs. Each of these six phonemes was represented in eight (or close to eight) different words: four times in an initial position and four times in a final position. In the four cases where the phoneme opened a word, it occurred twice in a stressed syllable and twice in an unstressed one. Also, in the four times where the phoneme occurred in the final position, it was twice in a stressed syllable and twice in an unstressed one. Part I of the table in Appendix 1 lists the forty-two words used in this test. Each word was written on a different card, in Hebrew script and with vocalization, accompanied by a sentence that helped to clarify its meaning.[15]

Ten native Hebrew informants of non-Oriental pronunciation who had completed high school were chosen. Before reading, each informant was asked to go over each card in order to ensure that s/he recognized all the words. The cards were shuffled randomly before each recording. The ten informants were recorded by a home cassette recorder. The length of these recordings was approximately fifteen minutes.

A list was formed, using the order of the words on the shuffled cards. Each word on a card was copied along with its corresponding word. For example, the third word pronounced by the first informant was /di'ra/ (דירה, 'flat'); the corresponding pair on the list was /ti'ra–di'ra/, written in Hebrew טירה / דירה ('castle' / 'flat'). On the whole, there were 420 pairs on the list (forty-two pairs for each of the ten informants).

Fifteen native Hebrew speakers (all of whom had finished at least high-school) listened to the 420 words. They were told that for each word heard there was a pair on the list. Their task was to circle the word they thought they heard.

1.1.2 Results of the First Perception Test

Whenever a participant circled a different word from the one the informant intended it to be when s/he read it, it was considered a "mistake."

---

15 The findings of this test were taken from a seminar paper done by Ms. Na'ama Bahat, under my tutelage, 1991. Exceptions to this ideal pattern were /p b/ and /k g/. [p b] are not common in final positions in modern Hebrew, because of the "frication law" in ancient Hebrew (O. Schwarzwald, *Grammar and Reality in the Hebrew Verb* [Heb.; Ramat-Gan, Israel: Bar-Ilan University Press, 1981], chapter 5). In the second perception test more Hebrew minimal pairs with final /p b/ were found. One pair of final [k g] was not recorded due to a technical problem.

There was almost complete harmony between the words read and the words perceived (circled). Out of the 6300 words heard in this test (42 words x 10 informants x 15 participants), only 104 were misidentified.

"Mistakes" in the perception of words could be attributed to a lack of concentration on the part of either the informants or the participants. Such lack of attention on the part of the informant could lead to slackness in speech, which will almost always cause a "mistake" by the listener. For example: One of the informants, whose words the listeners usually did not have any problem discriminating, pronounced one word, /ka'bai/ ('fireman' כבאי), in what seemed to be a careless way. As a result, ten of the fifteen participants did not perceive this word correctly. Except for such mistakes, most of the "errors" were random ones, and it seems that they occurred because of lack of concentration on the part of the listeners, which resulted either in reduced listening capacity or in erroneous marking. Perhaps the quality of the recordings, which were made by a home cassette recorder, also affected the perception to some degree.

As there is more than 98% correspondence between reading and perception, we can clearly conclude Hebrew speakers of the non-Oriental pronunciation make a physical differentiation between phonemes that are considered voiced and their voiceless counterparts.

1.2 Second Perception Test

1.2.1 Preparation of the Material and Recordings

Because the first test did not test the Oriental pronunciation of Hebrew, citation form is not the same as natural speech, and the words tested might not represent the various contexts in which these stops can occur in Hebrew, we added a second perception test, which was structured differently.

Each stop was represented by six common Hebrew words: in three at the beginning of a word and in the other three at the end. In order to verify that the adjacent vowel had no influence on the perception of the stop, we made sure that the chosen words contained front, back, and low vowels adjacent to each stop. This procedure was carried out for stops both in initial and final positions. The words were chosen in such a way that they could be used as minimal pairs. Part II of the table in Appendix 1 lists these words. We tried to choose pairs of words that were of equal frequency in contemporary Hebrew (lacking a scientific source, we depended on the collective intuition of myself and three Masters Degree students from our

Hebrew Language Department). Each word was embedded both in the beginning of a carrier sentence and at the end of such a sentence.

Altogether there were 108 utterances, fifty-four with stops in an initial position of the word and fifty-four with stops in a final position (6 stops X 3 different vowels X 3 positions = 54 utterances). Each was copied onto a card in Hebrew script and with vocalization. The cards were shuffled randomly and given to three native Hebrew informants of Oriental pronunciation, who had at least completed high school. Before reading, each informant was asked to go over each card to ensure that he recognized all the words. The recordings were carried out in a studio with professional recording equipment.

1.2.2 Results of the Second Perception Test

As in the first perception test, a list of 108 minimal pairs was prepared, and participants were asked to circle the word they perceived. Twenty-five native Hebrew speakers were chosen for the perception test (eight listened to the first informant, eight to the second informant, and nine to the third).

Almost 99% of the words were perceived correctly. Of the 2700 utterances heard by the twenty-five listeners (108 words X 25 listeners), only thirty-five words were identified erroneously.

1.3 Summary of the Two Perception Tests

In both tests we used minimal pairs, which included words that begin or end with a stop categorized as voiced or voiceless, and reached the same conclusions: Hebrew listeners have no problem in discriminating between words beginning or ending with phonemes categorized as voiced and voiceless stops (all else being equal). As a result, we must also conclude that Hebrew speakers pronounce voiced phonemes differently than their voiceless counterparts. Speakers obviously produce some physical difference that enables listeners to discriminate between utterances that include different phonemes. Our aim in the second stage of this study is to find a physical attribute that causes the voicing contrast in Hebrew.

2.0 Acoustical Analysis of Natural Hebrew Speech

Various languages contain shifts in acoustic features like duration of sounds, their fundamental frequency or "pitch," aspiration, etc., which accompany the physiological production of phonologically voiced-voiceless cognates. Some of these same shifts were found in Hebrew speech.[16] The second stage of this research sought to find a single distinct physical factor that by itself might cause the perceptual cue for the voiced-voiceless contrast in Hebrew speech. Lisker and Abramson have already shown that the simplest single measure in the acoustic signal which is acoustically efficacious in differentiating the voiced-voiceless contrasts in eleven languages is "voice timing."[17] In order to be able to compare data, we measured only "voice timing" at this stage of study.

Phonetic voicing is the physiological vibration of the vocal folds, and can be observed spectrographically through well defined acoustic characteristics. The obvious clue in a spectrogram for the vibration of the vocal folds is the "voice bar," which can be seen as vertical striations, at least on the bottom part of a "wide" spectrogram.[18] Even in ordinary waveforms it is easy to see and measure the voicing.[19] Following Lisker and Abramson's cross-language study of voicing, we assumed that the parameter which causes Hebrew listeners to differentiate between the categories we call

---

16 A. Laufer, *Synthesis by Rule of a Hebrew Idiolect* (Ph.D. thesis; University of London, 1972); A. Laufer, *Acoustic Phonetics* (Tel-Aviv, The Open University, 1984) units 5–7 of the course "Phonology of Modern Hebrew," pp. 37–38 (Heb.); A. Laufer, "'voiced / voiceless'"; L. J. Raphael, Y. Tobin, A. Faber, T. Most, B. H. Kollia, and D. Milstein, "Intermediate Values of Voice Onset Time," in *Producing Speech: Contemporary Issues*, eds. F. Bell-Bertti and L. J. Raphael [Katherine S. Harris, American Institute of Physics; New York: AIP Press, 1995), pp. 117–127.

For example, we found that Hebrew fits into universal tendencies with respect to segmental durations: We found that voiced sounds are shorter than their voiceless cognates, if pronounced in the same context and in the same conditions (compare, for example, to I. H. Slis and A. Cohen, "On the Complex Regulating the Voiced-Voiceless Distinction," *Language and Speech* 12 [1969] 80–102, 137–155; to I. Lehiste, *Suprasegmentals* [Cambridge, MA: MIT Press, 1970]; or to D. Byrd, "54,000 American Stops," in *UCLA Working Papers in Phonetics* 83 [1993] 97–115). Another acoustic feature which we found in Hebrew was that a vowel preceding a voiced sound is longer than the vowel preceding the voiceless sound, all else being equal. This tendency was found also in many other languages (e.g., P. A. Keating, "Universal Phonetics and the Organization of Grammars" in *Phonetic Linguistics: Essays in Honor of Peter Ladefoged*, ed. V. A. Fromkin [Orlando: Academic Press, 1985], pp. 115–132; A. C. Fowler, "Vowel Duration and Closure Duration in Voiced and Unvoiced Stops: There are No Contrast Effects Here," Haskins Laboratories Status Report on Speech Research, *Speech Research*, SR-107/108 [1991] 123–140). We also found a shift in fundamental frequency after voiceless sounds, in accordance with findings in other languages (e.g., O. Fujimura, "Remarks on Stop Consonants—Synthesis and Acoustic Cues," in *Phonetic and Linguistic Papers Presented to Eli Fischer-Jørgensen*, eds. L. L. Hammerich, R. Jakobson, and E. Zwirner [Copenhagen: Akademisk Forlag, 1971], pp. 221–232). If Hebrew speakers have adopted these acoustical differences to be perceptually relevant (that is, phonemic), this has yet to be tested.

17 L. Lisker and A. S. Abramson, "Cross-language Study."

18 E.g., see G. J. Borden and K. S. Harris, *Speech Science Primer: Physiology, Acoustics and Perception of Speech* (Baltimore: Williams and Wilkins, 1980), p. 178; A. Laufer, *Acoustic Phonetics*, p. 37.

19 E.g., P. Ladefoged, *Elements of Acoustic Phonetics*; P. Ladefoged, "Reading Waveforms" *Journal of the International Phonetic Association* 21 (1991) 32–35; J. E. Shope and L. L. Pfeifer, "Acoustic Characteristics of Speech Sound," in *Contemporary Issues in Experimental Phonetics*, ed. N. J. Lass (New York: Academic Press, 1976), pp. 171–224.

"voiced" and "voiceless" is the proportion of durations of the phonetic voicing.

2.0.1 Voice Timing

It is easy to compare measurements of voice timing in different sounds if we measure from a fixed point. Such a point can be the time of the release of the upper constriction of the vocal tract or the closure onset in the examined sound. The fixed point is customarily assigned the value "0," and time when voicing started or stopped is measured in relation to it. We will differentiate between sounds in initial and in final positions. For our research, medial sound can be treated either as initial or as final.

A. Sounds in an initial position. Lisker and Abramson coined the term Voice Onset Time (VOT) to describe the time difference between the onset of the vibrations of the vocal folds and the release of the upper articulators of a sound. Figure 1 demonstrates various values of VOT.[20] Part A depicts the condition of the upper articulators in two consecutive sounds, with the release of the first sound at time "0." The concurrent condition of the vocal folds is illustrated in part B, which includes three possibilities, depending on whether the vocal folds start vibrating before, at, or after the release. VOT has a negative value of -200 milliseconds on line 1, which means that the vocal folds start vibrating before the release of the sound. It has the value 0 on line 2, meaning that the vocal folds start vibrating exactly with the release of the sound. The positive value of +80 ms on line 3 means that the vocal folds start vibrating after the release of the sound.[21]

[20] Similar figures are used by modern phoneticians, e.g., P. Ladefoged, *Preliminaries to Linguistic Phonetics* (Chicago: University of Chicago Press, 1971), p. 10.

[21] This method of measuring VOT is widely accepted. Denis H. Klatt, however, defines VOT of voiced stops in a different way ("Voice Onset Time, Frication, and Aspiration in Word-initial Consonant Clusters" *Journal of Speech and Hearing Research* 18 [1975] 686–706 [reprinted in *Readings in Clinical Spectrography of Speech*, eds. J. Baken and R. G. Daniloff (Singular and Kay, 1991), pp. 226–246]).

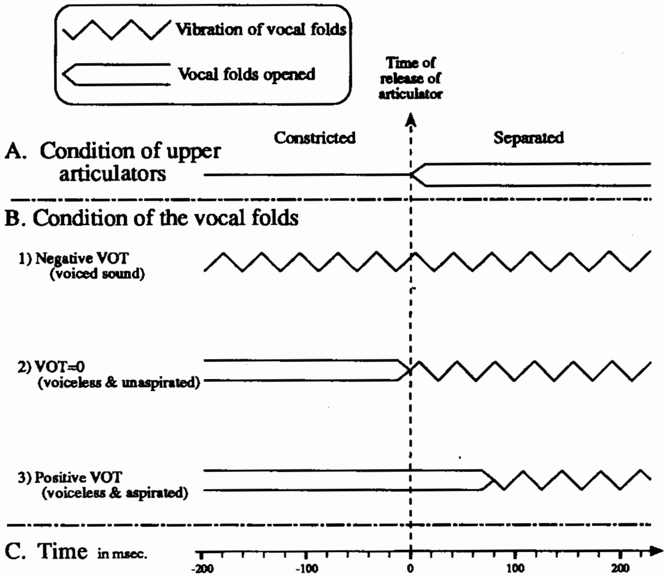

Figure 1. Possible relations between articulators and positions of the vocal folds in initial positions (VOT)

B. Sounds in a final position. For final positions, we measured Voice Timing in terms of Voice into Closure (ViC), an adaptation of the VOT method.[22] Here the point of reference is the onset of the constriction, which is assigned the time "0." Thus, ViC measures the time from the beginning of the constriction of the final sound until the cessation of voicing.[23] Figure 2, which is similar in structure to Figure 1, illustrates different measurements of ViC. On B3 the vibration of the vocal folds ceases 10 ms before the onset of the closure, and therefore ViC = –10 ms. On B2, ViC = 100 ms, meaning that the vibration of the vocal folds ceases simultaneously with the release of the consonant. (In this figure the time release of the articula-

---

[22] Similar measurements of what we call ViC were used for instance by M. Fourakis and K. G. Iverson, "On the 'incomplete neutralization' of German Final Obstruents," *Phonetica* 41 (1984) 141 and by R. F. Port and M. L. O'Dell, "Neutralization of Syllable-final Voicing in German" *Journal of Phonetics* 13 (1985) 455–471.

[23] We defined the time "0" at the cessation or at the drastic weakening of the second formant of the preceding vowel. This was also the criterion of M. Fourakis and K. G. Iverson, "On the 'incomplete neutralization' of German Final Obstruents," p. 143. For determining the time "0," we don't take F1 into account: Sometimes the amplitude of F1 is still strong at this point, and sometimes it is weakened drastically before this point. The last phenomenon is known as F1 cutback (e.g., A. M. Liberman, P. C. Delattre, and F. S. Cooper, "Some Cues for the Distinction Between Voiced and Voiceless Stops in Initial Positions," *Language and Speech* 1 [1958] 153–167).

tor was arbitrarily assigned to 100 ms.) B1 illustrates ViC = +170 ms, which means that the vibration of the vocal folds extends beyond the release of the upper articulator; this phenomenon is common in Hebrew, where the voicing of absolute-final voiced stops is often prolonged into the release. Phonetically, this prolongation can be considered a short vowel, but Hebrew listeners perceive it as part of a final stop (a release), not as a new vowel. Since we are interested in comparing the voice timing of voiced and voiceless stops, we include this "post-closure voicing" as part of the voicing of the final sound.[24]

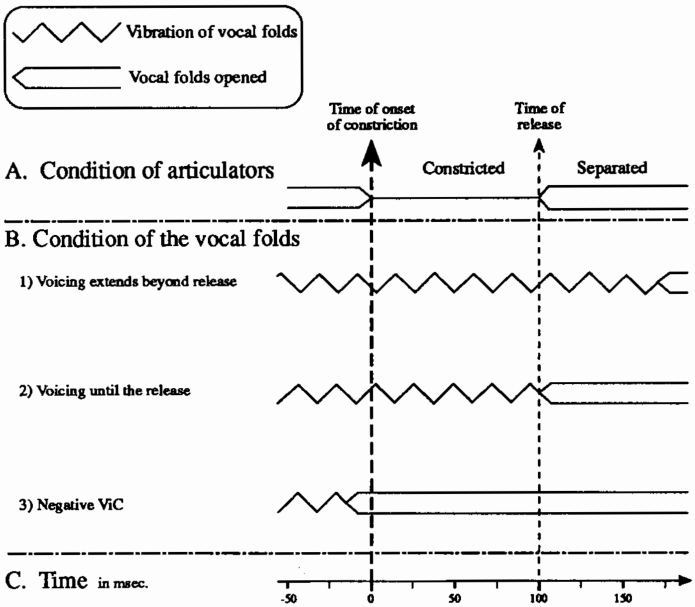

Figure 2. Possible relations between articulators and positions of the vocal folds in final positions (ViC)

Both VOT and ViC measure voice timing. We expect that voicing will start earlier in an initial voiced sound than in its voiceless counterpart, that

[24] The phenomenon of 'post closure voicing' was common in our recordings of Hebrew, and here we will only sum up the data of our three Oriental speakers: In absolute-final voiced stops extension of voicing beyond the beginning of release occurred 37 times out of 54 possible occurrences. Two of the three Oriental speakers had this phenomenon as a dominant characteristic, and the third produced it rarely (5 times out of 18 possible occurrences). In this corpus the shortest 'post closure voicing' we found was of a single striation, and the longest prolonged to 110 ms after the beginning of release. We also found that when Hebrew speakers want to imitate French pronunciation, one of the characteristics is the exaggeration of this 'post closure voicing.'

is, that the VOT of the voiced sound will have a greater negative value than the VOT of the voiceless counterpart. On the other hand, we would expect the vocal fold vibration to persist longer in a voiced final sound than in its voiceless counterpart, that is, that the ViC of the voiced sound will have a higher value than the voiceless counterpart.

C. Sounds in medial position. Because we can usually see both the release and the onset of constriction in medial position sounds, we can measure the voice timing with either the VOT or the ViC method.[25]

2.1 Measurements of Voice Timing in Hebrew

We measured the voice timing of stops using three separate sets of recordings. The first was made under studio conditions by a single native male speaker of Hebrew with a non-Oriental pronunciation, reading isolated sentences and speaking spontaneously in an ordinary rhythm of speech—neither slow nor fast. The other two sets were the structured recordings mentioned in the previous section. The recordings were usually measured twice, by two different people. If the two measurements were close, an average was taken. On the rare occasions when the measurements were wide apart, a third measurement was taken, and an average was drawn between the two closest measurements.

The first recording lasted about five minutes and was analyzed by a Kay spectrograph. From wide band spectrograms we measured the voice timing of the /p b t d k g/ in this corpus. We measured the VOT of 170 stops including forty-five in initial positions and 125 in medial positions. Twenty-four stops occurred in final positions, for which we measured their ViC.

These measurements showed clearly that voice timing can distinguish voiced and voiceless consonants in Hebrew, with voiced sounds always having a considerably longer time of vocal fold vibrations than voiceless sounds. Although this recording could be considered ideal from the points of view of quality and naturalness, especially since this corpus was not structured for examining voicing, it has two disadvantages: 1) The number of occurrences of the different stops was not uniform; for instance, there

---

[25] In absolute-final voiced stops, as noted above, voicing often prolongs into the release, and our ViC measurements include this 'post closure voicing' as part of the voicing of the final sound. In medial voiced sounds followed by another voiced sound the voicing in our ViC measurements was measured only to the end point of the visible fricative release of the stop (beyond this release we assumed that the next voiced sound begins). We found that in Hebrew the release of such non-absolute-final voiced stops is much shorter than in absolute-final position. Anyhow, for medial stops we usually measured the VOT in this study, except for part of our third set of recordings (we will come back to our three sets later).

was not even one example of a final /g/ /p/ or /b/, but nineteen of final /t/. 2) It would have been preferable to have a greater number of informants. To address these issues, we added measurements from the two sets of recordings used in the first and second perception tests mentioned earlier. These were analyzed with the "Signalyze" program.[26] Waveforms and synchronized wide band spectrograms were taken, which helped in defining the borders in which we were interested, and the time measurements were given automatically by the program.

As mentioned above, the second set of recordings was the speech of ten non-Oriental informants (five females and five males), pronouncing Hebrew words that have stops in initial and in final positions. In this set we measured 240 stops in an initial position and 160 in a final position; half of these were voiced and half voiceless. The third set of recordings was the speech of three Oriental male informants, in which we measured 324 stops (162 voiced and 162 voiceless) in initial, medial, and final positions.

Discussion and results. Pulses more or less at the frequencies of the surrounding vowels during the closure were treated as voicing, since they were undoubtedly the result of glottal activity.[27] We considered these glottal pulses as "voice" without paying attention to their intensity, even when they were faint.

In all voiced stops we saw low-frequency vertical striations (voice-bar) during most of the closure period on the spectrograms. In most cases we saw the voice-bar for the entire duration of the stop; where the closure duration was relatively long, we could see a decrease in the amplitude of the voice-bar towards the release. About 10% of the time we could not see any glottal activity just before the release, but normal voicing resumed immediately after the release. In such stops the oral pressure builds up to the extent that it causes the cessation of the vibration of the vocal folds. After the release the pressure drops down and the vibration of folds can be resumed.[28] In all such sounds we assume that the vocal folds are in the posi-

---

[26] We used "Signalyze"—*Signal Analysis for Speech and Sound,* InfoSignal, Lausanne, Switzerland, (1994) version 3.0, Erich Keller. In "Signalyze," the wide band spectrogram is time synchronized with the waveform and it is thus comparatively easy to measure voice timing.

[27] L. Lisker and A. S. Abramson, "Cross-language Study," n. 14.

[28] D. H. Klatt, "Voice Onset Time," p. 699; J. C. Catford, *Fundamental Problems,* Chapter 3; J. M. Pickett, *The Sounds of Speech Communication* (Baltimore: University Park Press, 1980), pp. 111–112, 123 (3 chapters of the book are reprinted in: *Readings in Clinical Spectrography of Speech,* eds. J. Baken and R. G. Daniloff [Singular and Kay, 1991], pp. 75–123); A. Laufer, "Does the 'voiced epiglottal plosive' Exist?" *Journal of the International Phonetic Association* 21 (1991) 44.

tion of voicing, which may weaken or even stop because of aerodynamic forces. So, in addition to not taking the intensity of the glottal pulses into consideration, in sounds that had a few striations missing before the release, we treated this part of the sound as "voiced," without any trace of glottal pulsation. We did this because we do not have any statistical measure to describe a relation between the intensity of glottal pulses and their perceptual correlation and because voice timing underlies a complex set of interrelated acoustic features any one of which may have perceptual efficacy.

Our data show individual differences as well as contextual ones. The measurements point to systematic differences that are affected by place of articulation and vocalic environment. In agreement with the general tendency noted by numerous scholars,[29] we found that voiceless velars have longer VOT than labials or alveolars in Hebrew and that voiceless Hebrew alveolars usually have a slightly longer VOT than labials. This is in accordance with other findings in Hebrew[30] and in English.[31] Our data also show that Hebrew speakers tend to have longer VOTs in voiceless stops in the vicinity of high vowels, especially /i/, in accordance with Klatt's findings in English[32] and Nearey's and Rochet's findings for French and English speakers.[33]

However, all these differences of voice timing can be considered of secondary importance in contrast to the difference of voice timing between the "voiced / voiceless" categories. From the perspective of voice timing, there was no fundamental difference between a stop in a single word and the same stop in the same word when it was part of the beginning or the end of a longer utterance. Nor could we find statistically significant variances in VOTs or in ViCs of stops adjacent to a front, back, or low vowel when compared to their voiceless cognates.[34] In accordance with Lisker &

---

[29] E.g., L. Lisker and A. S. Abramson, "Cross-language Study," p. 399; D. H. Klatt, "Voice Onset Time," pp. 686, 692; P. A. Keating, "Phonetic and Phonological Representation of Stop Consonant Voicing" *Language* 60 (1984) 286–319; T. M. Nearey and B. L. Rochet, "Effects of Place of Articulation and Vowel Context on VOT Production and Perception for French and English Stops" *Journal of the International Phonetic Association* 24 (1994) 1–19; D. Byrd, "54,000 American Stops," p. 102; L. Obler, "The Parsimonious Bilingual," in *Exceptional Language and Linguistics*, eds. L. Obler and L. Menn (New York: Academic Press, 1982), pp. 339–346; L. J. Raphael *et al.*, "Intermediate Values."

[30] L. Obler, "Parsimonious Bilingual," p. 342, and L. J. Raphael *et al.*, "Intermediate Values."

[31] For example, see D. H. Klatt, "Voice Onset Time," pp. 692, 695, or D. Byrd, "54,000 American stops."

[32] D. H. Klatt, "Voice Onset Time," p. 694.

[33] T. M. Nearey and B. L. Rochet, "Effects of Place of Articulation."

[34] These findings are in accordance with Klatt's findings in English ("Voice Onset Time," pp. 691, 694).

Abramson and with Klatt,[35] our data also revealed that the word data agree by and large with the sentence data and the spontaneous speech data. As our prime interest at this stage is to find an acoustical basis for the phonological distinction of voicing, we can disregard all these secondary small differences and group all our measurements for each phoneme together.

Our findings showed that the difference in voice timing between a voiced sound and its voiceless counterpart is very similar for all our informants, whether of Oriental or non-Oriental pronunciation. From the perspective of voice timing, there is no statistically significant difference between all our fourteen Hebrew speakers; we could, therefore, classify all our measurements in one group.

2.2 Summary of the Acoustical Findings of the Three Recordings

Our three sets of recordings complement each other. Since personal differences were negligible from the voice timing perspective, the data of our fourteen Hebrew informants can be summed up into two tables (Appendix 2).

These data are transposed into two figures: the average VOT according to the data in Part I of Appendix 2 (Figure 3) and the average ViC according to the data in Part II of Appendix 2 (Figure 4).[36]

---

[35] L. Lisker and A. S. Abramson, "Cross-language Study," p. 413; D. H. Klatt, "Voice Onset Time," pp. 672, 691.

[36] Obler's report of VOT of Hebrew unilinguals, "Parsimonious Bilingual," p. 342, is in general agreement with our data. Our measurements of Hebrew voice timing was first published in A. Laufer, "Voicing." We find recently supportive evidence for our VOT measurements from the data published by L. J. Raphael *et al.*, "Intermediate Values." They published VOT data of 23 Israeli subjects reading 15 minimal pairs of Hebrew words beginning with the same 3 sets of stops we did. Their VOT results are very similar to ours. L. J. Raphael *et al.* did not measure ViC in Hebrew, so we could not compare results. (I would like to thank Alice Faber, from Haskins Labs., for sending me L. J. Raphael *et al.*'s paper.)

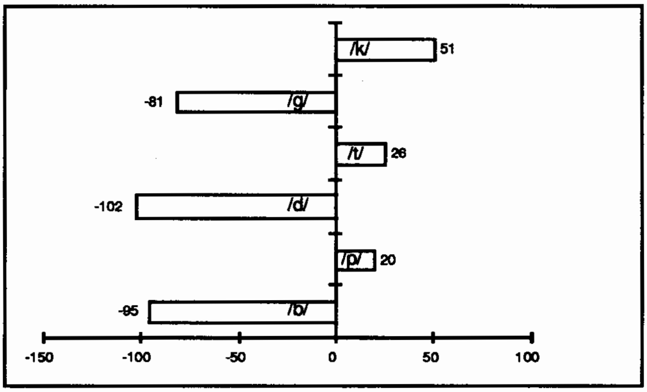

Figure 3. Hebrew averages of VOT in ms (14 Speakers–data of each speaker were based on at least 69 measurements.)

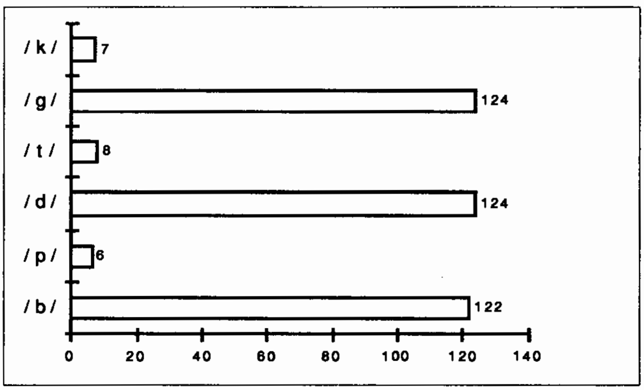

Figure 4. Hebrew averages of ViC in ms (14 Speakers–data of each speaker were based on at least 69 measurements.)

Our findings showed that the onset of phonetic voicing (VOT) or its cessation (ViC) can be used to distinguish between the "voiced stops" and their "voiceless" counterparts. For each of our informants the voice timing distributions of /b d g/ and /p t k/ occupy non-overlapping ranges. Figure 3 shows clearly that for non-final voiced stops voicing starts well in advance of the release, while voicing of voiceless stops always starts after the release. Figure 4 shows that the voicing of the voiced stops extends much longer into the closure than that of their voiceless cognates. The analysis of each individual subject shows clearly that the phonetic voicing in the production of Hebrew voiced stops continues for a considerably longer period of time than in their voiceless counterparts.

Although our data show that the VOT of /k/ is longer than that of /t/ or /p/, we have already pointed out that such differences are of secondary importance in comparison to the difference found between the "voiced / voiceless" categories. As far as voice timing is concerned, the noticeable aspiration of /k/ is not significantly different, so we can group all the voiceless stops together, as we did at the bottom of parts I and II of Appendix 2, in order to compare our data in Hebrew to published data for other languages.

2.3 Comparing Voice Timings of Hebrew to VOT of Other Languages

We can compare our findings with those for other languages. Lisker and Abramson checked VOTs in eleven languages, in all of which they found a time difference in the onset of voicing which could differentiate between the group of voiced stops and their voiceless counterparts.[37] Figure 5 is constructed from their findings.[38] To this, we have added Keating's measurements of VOT in Polish[39] as well as our findings for contemporary Hebrew. The horizontal axis in Figure 5 represents the VOT in milliseconds, with dots placed to show the average of the VOT of stops in an initial position. The dots represent the categories in each language: If there are two categories of stops (voiced and voiceless) there are only two dots, the left one representing the average for voiced stops and the right dot the average for voiceless stops. (If a language has three categories of stops, the middle dot represents the average of voiceless unaspirated stops, and the right dot that of voiceless aspirated stops. For English there were two discontinuous ranges of measurements for the voiced stops, and therefore Lisker and Abramson enclosed the two left dots for English in a box to show the free variation of prevoicing versus short lag. The two bottom languages have four categories of stops, such as /b, b^h, p, p^h/.[40] (The voiced aspirated stop in these languages is represented by an asterisk in Figure 5.)

---

[37] L. Lisker and A. S. Abramson, "Cross-language Study."

[38] L. Lisker and A. S. Abramson, "Distinctive Features," p. 772.

[39] P. A. Keating, "Phonetic and Phonological Representation," figure 1, p. 300.

[40] E.g., P. Ladefoged, *Preliminaries to Linguistic Phonetics*, pp. 13–14.

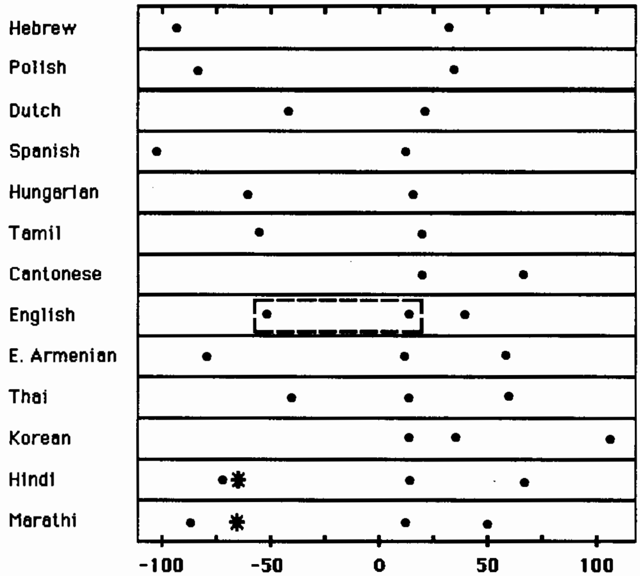

Figure 5. VOT averages of initial stops in 13 languages

This figure[41] suggests that each language divides up the voice timing continuum differently. For every language, the phonemic "voice" features are realized phonetically in relative terms of "more" and "less" voicing. Moreover, the range between categories (the distance between two dots) has different values in different languages, just as the borders between categories are in different places along the VOT scale. We can conclude from this that from the VOT perspective one phonetic sound can represent a voiced sound in one language and a voiceless sound in another.[42]

---

[41] The figure is based on the findings of L. Lisker and A. S. Abramson of eleven languages. We calculated Keating's data (Keating, "Phonetic," figure 1, p. 300) on Polish VOT and added it on top of the figure, along with our findings in Hebrew. Release of stops is in time 0. The left dot in each language represents the VOT average of the voiced stops. In languages that have only two dots the right dot represents the average of the voiceless stops. In languages that are marked with three dots the middle dot represents the voiceless unaspirated stops, and the right dot represents the average voiceless aspirated stops. This figure was adapted from L. Lisker and A. S. Abramson ("Distinctive Features," p. 772.), and I wish to thank the authors, the Linguistic Society of America and *LANGUAGE* for their permission to use *the Figure*.

[42] By the way, these physical differences of VOTs may lead us to a perceptual assumption: the same phonetic sound can be perceived in one language as a voiced sound and in another language as a voiceless sound. This was verified experimentally (e.g., J. Lotz, A. S. Abramson, L. J. Gerstman, F. Ingemann, and W. J. Nemser, "The Perception of English Stops by Speakers of English, Spanish, Hungarian and Thai: A Tape-cutting Experiment" *Language and Speech* 3 [1960] 71–77; J. E. Flege and W. Efting, "Imitation of a VOT Continuum by Native Speakers of English and Spanish: Evidence for Phonetic Category Formation"

Figure 5 also shows that the data for Hebrew are similar to those for Polish and Spanish in these respects: The range between the categories is wide and tends towards the left on the VOT scale (which means long voicing lead of voiced stops). At the same time, Hebrew, Polish, and Spanish differ from Cantonese and English in that the range between the categories in Cantonese is narrow and tends towards the right on the VOT scale. In spite of the similarities, there are also differences: For example, Spanish voiceless stops are usually produced with minimal positive VOT values (minimal aspiration), while in Hebrew and in Polish the VOT of voiceless stops is longer.[43]

We were unable to find ViC studies in other languages with which we could compare our Hebrew findings. However, the data in Figure 4 show that the range of Voice Timing in final stops is quite wide, just like the VOT for Hebrew stops. Voicing ceases much later in Hebrew speakers' final voiced stops than in their voiceless counterparts.

Although our data definitely support the view that voice timing is an effective measure for differentiating homorganic stop categories, it remains to be demonstrated that this feature is perceptually relevant and sufficient. It could be that another physical factor which we did not measure or note (e.g., force, air pressure, aspiration, muscle tension, duration of the consonant, duration of vowels in the vicinity of the consonant, pitch, F1 transition) is the one that allows Hebrew speakers to differentiate between the two categories. It could be argued that voice timing is only an accompanying feature. Therefore, the following section is devoted to testing if voice timing is a sufficient differentiator for perceiving voiced / voiceless categories.

The Journal of Acoustical Society of America 83 [1988] 729–740; A. S. Abramson, "Laryngeal control in the plosive of standard Thai" *PASAA: A Journal of Language Teaching and Learning in Thailand* 19 [1989] 85–93).

[43] In other terms we can say that both Hebrew and Polish have a long voicing lead on the one hand and an intermediate lag on the other hand (compare to L. J. Raphael *et al* "Intermediate Values," p. 118). One possible suggestion is that the similarities of voice timing of Hebrew, Polish, and Spanish can be attributed to Hebrew having been "revived" as a spoken language by speakers of the dominant substratum languages of Polish and Spanish. (This interesting hypothesis was raised by one of my anonymous reviewers, and I would like to thank him for it.) This is possible, although English was and is in the background of Israeli Hebrew (during the British Mandate, and under heavy American culture influence), and yet Hebrew and English are very different in their voice-timing rules. Anyhow, we are not equipped to answer this interesting question yet, and it is not our goal in this study.

3.0 Perceptual Tests of Voice Timing of Hebrew Speakers

The finding that the physical voicing of a voiced stop is always longer than its voiceless counterpart does not alone prove that voice timing is the perceptual cue for the voice category for two main reasons: (1) Investigators have found acoustical features that do not influence speech perception;[44] (2) Tene, Chayen, and Chayen and Dror have argued that voicing is only an accompanying feature in Hebrew and not a phonemic one; they believe that the essential characteristic in Hebrew is the force of articulation.[45] This last argument has never been experimentally proven for Hebrew (or for any language for that matter).[46] In this section we want to prove that in Hebrew the voicing feature is perceptually relevant and that it is a sufficient cue. The best way to do this is with an identification psychoacoustic test, in which subjects are presented with a set of randomized stimuli that differ only in the feature that is being investigated and must classify each stimulus according to a set of labels provided by the tester.[47] Accordingly, we designed a perceptual identification test by using synthetic stimuli on a voiced / voiceless continuum that differ **only** in voice timing.

3.1 Preparations for the Identification Psychoacoustic Tests

3.1.1 Synthesis of the Series of Utterances for Testing Voicing

The synthetic speech was produced by a synthesizer that was adapted to the acoustic findings of Hebrew speech.[48] It "speaks" as a result of the

---

[44] See for example L. Lisker and A. S. Abramson, "Cross-language Study," p. 416; R. F. Port and P. Crawford, "Incomplete Neutralization and Pragmatics in German" *Journal of Phonetics* 17 (1989) 274.

[45] D. Tene, "Measured Duration," §§ 4.6, 5.3; M. J. Chayen, "Pronunciation of Israeli Hebrew" pp. 288, 290; M. J. Chayen and Z. Dror, *Introduction to Hebrew*, p. 294.

[46] There are many works showing that a range of acoustic or physiological features can be used to discriminate phonologically between voiced and voiceless cognates. These features are connected to aspiration, duration of the consonant, duration of vowels in the vicinity of the consonant, F1 transitions, and fundamental frequency (see for example O. Fujimura, "Remarks on Stop Consonants;" L. Lisker and A. S. Abramson, "Cross-language Study" and "Distinctive Features"; A. S. Abramson and L. Lisker, "Relative Power of Cues: F0 Shift Versus Voice Timing" in *Phonetic Linguistics: Essays in Honor of Peter Ladefoged*, ed. V. A. Fromkin [Orlando: Academic Press, 1985], pp. 25–33.; D. H. Whalen, A. S. Abramson, L. Lisker, and M. Mody, "Gradient Effects of Fundamental Frequency on Stop Consonant Voicing Judgments" *Phonetica* 47 [1990] 36–49; A. Laufer, "'voiced / voiceless'"). However, the features that are found to be perceptually relevant are the consequence of the laryngeal activity and are included in what we call here "voice timing" (A. S. Abramson, "Laryngeal control"; A. Laufer, "'voiced / voiceless'"). The fact is that no linguist or phonetician has proven the existence of a physical or physiological substance that we call here "force," and that serves as a cue for the voice-voiceless contrast.

[47] A. M. Liberman *et al.*, "Cues for the Distinction."

[48] We used an LSI Research Speech Synthesizer, which is a parallel formant one, based on Holmes' synthesizer (D. J. Quarmby and J. N. Holmes, "Implementation of a Parallel-Formant Speech Synthesizer

computer supplying it, every 10 milliseconds, with a "frame" that contains values for its twelve parameters. The production of the computer software can be thought of as a table-like file, in which each row represents a frame and their sequence represents the progression of time. The columns are the twelve parameters, representing primarily the fundamental frequency, formants frequencies, formants amplitudes, and degree of voicing. We can arrange for the computer to give different values to the parameters in each frame.

The "basic file" contained appropriate values of the twelve parameters for producing a sound that was perceived as the Hebrew word /bi/ ('in me'). In this synthetic word, the stop duration of /b/ was 150 ms, the release was at the sixteenth frame, and the following vowel lasted 270 ms.

As we wanted to verify our assumption that the voice timing cue is responsible for differentiations between voiced and voiceless counterparts, we produced synthetic words that were identical except for the values representing the degree of voicing. In order to do this, we duplicated our "basic file" for producing the Hebrew word /bi/ thirty-one times. In each one of these copied files, we changed only the parameter that signified the degree of voicing, leaving the values of the other eleven parameters exactly the same. At the end of this process, we had a series of thirty-one files for producing thirty-one synthetic utterances that differed only in their VOT values.

In this experiment we used only two values for the parameter of voice excitation. These were the two extremes: completely voiceless excitation (hiss) or completely voiced excitation (buzz). We altered the voice parameter of the thirty-one files gradually: In the first file, the first frame was given the value for "hiss" (voiceless) and all the rest of the frames in the utterance were given a value for "buzz" (voiced). In the second file, the first two frames were given the voiceless values and all the rest of the frames were given the value for "voiced." In such a gradual fashion we changed all the thirty-one files, so that in the thirty-first file the first thirty-one frames were given the value "voiceless" and all the rest of the frames were given the value "voiced."

Using a Single-chip Programmable Signal Processor," *IEEE Proceedings* 131 [1984] 563–569). Using "Syncon" software (J. N. Holmes, *Syncon—A synthesis by rule software package for convenient interactive control of the Loughborough Sound Image speech synthesizer using a BBC microcomputer*, Millgrant Wells, Rugby CV21 3UF, England [1986]), we gave acoustic values based on Hebrew speech, and the result was that the Synthesizer produced utterances that were perceived as Hebrew speech.

In the sixteenth frame of each file the values of all the parameters produced the release of the stop; therefore, we could assign the value "0" to the sixteenth frame and describe the aforementioned process in terms of VOT. As a result, we got a series of thirty-one synthetic utterances, initiated by a stop and differentiated from each other only in their VOT values. In this way, VOTs were increased by steps of 10 ms until the thirty-first file. The first file produced an utterance initiated by a stop with VOT of -150 ms, the second file produced an utterance initiated by a stop with VOT of -140 ms, etc. The sixteenth file produced an utterance of 0 VOT, in the seventeenth file the VOT was +10...and in the thirty-first file the VOT was +150 ms. The consequence of this process was a series of thirty-one identical stimuli with initial stops that differed from one another only in the voice timing; these thirty-one utterances ranged from a stimulus the voicing of which starts 150 ms before release to a stimulus the voicing of which starts 150 ms after release.[49]

We randomized the order of the files, synthesized each of them twice, and recorded them on tape. Each pair of synthetic utterances was labeled, with natural voice, by ordered numbers. This recording was called "the /bi-pi/ series." It is worthwhile to reiterate that the thirty-one synthesized words were exactly the same except for their VOT values.

We prepared two additional series in an identical fashion: "the /dam-tam/ series" and "the /gam-kam/ series." As a result, we had three series of synthetic words with initial stops that differed only in their VOT values: from -150 ms to +150 ms, while the progression of each utterance was by 10 ms. Our goal was to find out if voice timing by itself is a sufficient differentiator of the voiced / voiceless categories in Hebrew. Since the principal difficulty in voicing results essentially from stops in initial positions,[50] we restricted our perception tests to initial stops.

3.1.2 The Response Sheet

We prepared a two page response sheet, written in Hebrew, including three parts. Each part had thirty-one iterations of a minimal pair that is meaningful in Hebrew: In the first part "bi-pi" appeared thirty-one times,

---

[49] As in our system the parameter of voicing is controlling the voiced or the voiceless excitations, the result was that for every positive VOT an aspiration was produced for the duration of the VOT (hiss excited the formants). When the voiceless excitation was "on" in the closure period, it did not have any significant aspiration. (That is because in the stop period all the amplitudes were shut off, except for F1 which had a low amplitude.)

[50] E.g., A. Laufer, "'voiced / voiceless.'"

in the second part "dam-tam"; and in the third part "gam-kam."[51] We also included in the response sheet some questions that might have some bearing on the results (e.g., age, sex, and knowledge of other languages).

In the introduction to the tape, participants were asked to listen carefully to the recording and to decide which word of the respective pairs resembled the utterance they heard most closely. We made the subjects aware that each utterance would be heard twice and that it might not necessarily sound natural; they were told to circle the word that they felt the utterance mostly resembled (forced choice between two).

3.1.3 The Psychoacoustic Identification Tests

About a third of the subjects listened to the recordings through high-quality head-sets; the rest listened through an ordinary cassette recorder in an ordinary room. Each test lasted about twenty minutes and included two short breaks.

3.2 Results of the Perception Tests

We tested 207 subjects, but did not take into consideration incomplete response sheets or those from subjects who declared that Hebrew was not their main language. As a result, only 190 response sheets were fit to be used for the purpose of our study.

There were no significant differences between the "listening condition groups" – not between subjects that listened to the recordings using head-sets and those who listened through an ordinary tape recorder, not between males and females, not between subjects of different ages, nor between subjects who spoke only Hebrew and those who spoke more languages. We were, therefore, able to pool the 190 response sheets into a single group.

3.2.1 Results of the /bi-pi/ Series

Figure 6 summarizes the perceptions of the /bi-pi/ series according to the 190 response sheets. The horizontal axis represents the voice timing of the thirty-one utterances that composed the series, ordered according to their VOT values from -150 to +150 ms. As mentioned, all the stimuli were identical except for the parameter that is responsible for the VOT changes; therefore, the horizontal axis can symbolize the thirty-one syn-

---

[51] In the "bi-pi series" /bi/ means 'in me,' and /pi/–'my mouth'; in the "dam-tam series" /dam/ means 'blood,' and /tam/–'finished,' or 'innocent'; in the "gam-kam series" /gam/ means 'also,' and /kam/–'wakes up.' According to my intuition, both words in every series have a similar frequency in colloquial Hebrew.

thetic words. It is worth emphasizing that although the values have been arranged in an ascending sequence, they were heard by the subjects in random order. The vertical axis indicates the percentage of respondents who identified the sound as /bi/ (a filled square) or /pi/ (an empty square).

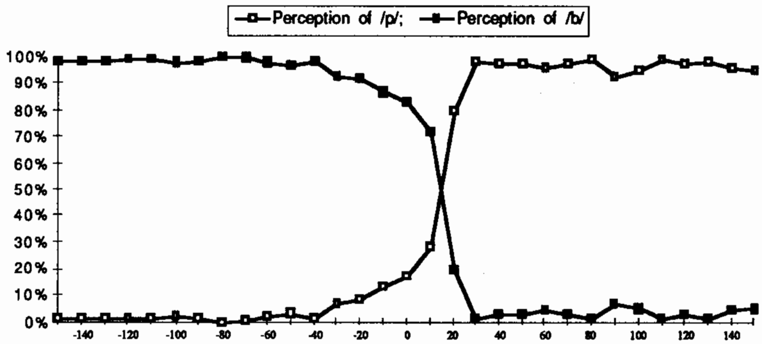

Figure 6. Identification of /bi-pi/ as a function of VOT by 190 Hebrew subjects

Had there not been a strong tie between the changing values of VOT and the categorical perception of voicing, we would have expected to find irregularities in the graph. Figure 6 clearly shows a strong tie between VOT values and the perception of voicing: When VOT was between -20 and -150 ms, more than 90% of the participants perceived /b/; between +30 and 150 ms, more than 90% perceived /p/. Only in four cases was the percentage of perception under 90%—where VOT was -10, 0, +10 and +20 ms. If we treat 70% agreement as sufficient for relative certainty, then the dividing line between /bi/ and /pi/ for most subjects is VOT = +10 ms. The majority of subjects identified /bi/ if VOT was less than +10 ms, and identified /pi/ if VOT was greater than +10.

3.2.2 Results of the /dam-tam/ Series

Figure 7, which is constructed in the same fashion as Figure 6, represents the perceptions of /dam-tam/ as a function of voice timing by the same 190 Hebrew subjects. These graphs show a greater uniformity than the graphs of the /bi-pi/ series. Here, more than 95% of the participants perceived the voiced /d/ when VOT values were between 0 and -150 ms, while more than 95% of the participants perceived the voiceless /t/ when VOT values were between 30 and +150 ms. In this series, the percentage of identification was less than 95% in a more narrow area than in the /bi-pi/ series. The dividing line for most subjects was VOT = +15 ms. The majority of subjects (more than 80%) perceived /dam/ if VOT was less than +15 ms, and /tam/ if VOT was more than +15 ms.

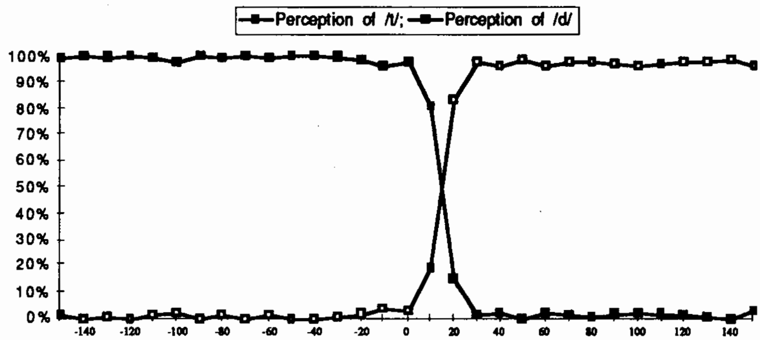

Figure 7. Identification of /dam-tam/ as a function of VOT by 190 Hebrew subjects

3.2.3 Results of the /gam-kam/ Series

Figure 8 presents the perception of /gam-kam/ by the same 190 subjects. Here too the graphs show a clear tie between the VOT values and the categorical perception of voicing. In this figure there is not a sharp dividing line that is applicable to the majority of subjects; there is a dividing zone that lies between +10 and +20 ms. When VOT was less than +10 ms, more than 84% of our subjects identified the voiced /gam/; when VOT was greater than +20, more than 84% identified the voiceless /kam/.[52]

[52] The graphs for /gam-kam/ in Figure 8 are less uniform than those of /dam-tam/, especially in the range between +30 and +150 ms of VOT. In this range the graphs are not completely horizontal: they fluctuate in a zone of 16%. In addition, we can see that the categorical boundary between /g/ and /k/ is wider than the other two boundaries. The reasons for these may be the following: 1) The initial synthesized stimulus in the /gam-kam/ series did not sound natural enough; 2) it is feasible that more "mistakes" were made due to a lack of alertness that was caused by the fact that the /gam-kam/ series was the third and last test. Nevertheless, even these graphs show a very clear tie between VOT values and the categorical perception of voicing.

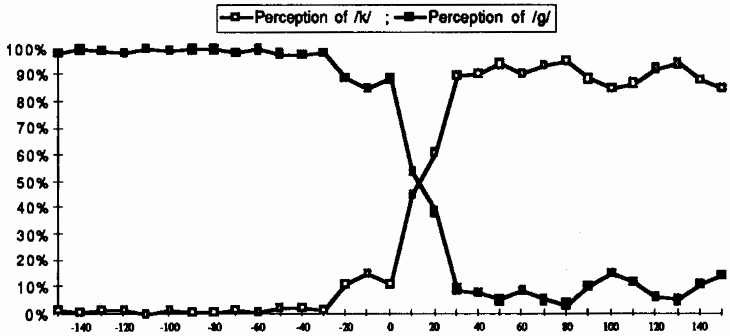

Figure 8. Identification of /gam-kam/ as a function of VOT by 190 Hebrew subjects

3.3 Comparing the Identification of the Voicing Contrast in Hebrew and in Other Languages

Abramson and Lisker carried out similar perception tests of synthetic speech for English, Spanish, and Thai listeners.[53] As expected, their Thai listeners divided the VOT continuum into three distinct categories in the vicinities of their phoneme boundaries. On the whole, our Hebrew results are similar to their graphs for English and Spanish listeners, which also show that maximum identification of a voiced sound occurs when its VOT is in the range from a very negative value to a low positive value (immediately after the release), whereas the perception of a voiced sound strives to a minimum when the value of VOT ranges from a low positive VOT to a very high positive VOT. In this respect there are no substantial differences between the identification of Hebrew listeners and that of English or Spanish listeners. All these data of the synthetic stimuli provide clear enough cues for two (or three) good perceptual categories at each place of articulation.

Abramson and Lisker listed the points at which more than 50% of the listeners begin to perceive a voiceless sound instead of a voiced one for their English and Spanish subjects.[54] We have added to it our Hebrew data in Table 1. After each value we included, in parentheses, some data of the

---

[53] A. S. Abramson and L. Lisker, "Voice-timing Perception on Spanish Word-initial Stops" *Journal of Phonetics* 1 (1973) 1–8; A. S. Abramson and L. Lisker, "Relative Power of Cues: F0 Shift Versus Voice Timing."

[54] A. S. Abramson and L. Lisker, "Voice-timing Perception," p. 4.

production of the voiceless stops: the mean VOT and in square brackets the range between the minimal and the maximal VOT.[55]

Table 1. Category boundaries in perception of voice timing (ms)

	Hebrew	Spanish	English
Labial [p b]	**+14** (M: 20; [0-59])	**+14** (M: 4; [0-15])	**+25** (M: 58; [20-120])
Apical [t d]	**+15** (M: 26; [0-66])	**+22** (M: 9; [0-15])	**+35** (M: 70; [30-105])
Velar [k g]	**+13** (M: 51; [17-107])	**+24** (M: 29; [15-55])	**+42** (M: 80; [50-135])

Note that the Hebrew 50% perceptual crossovers have the lowest VOT values in the vicinity of +14 ms, while for English listeners the perception crossover point always has a greater value and is different for each place of articulation. It is tempting to say that the higher positive value of the perception crossover point for English listeners is due to the fact that English speakers aspirate the voiceless sounds more than Hebrew and Spanish speakers, but the link between performance and perception is not that simple.

Every speaker has two phonological systems: one for performance and another for perception.[56] With the help of the performance phonological system, the speaker realizes the abstract units that compose the linguistic level in the brain into a physical sound wave. This system is very flexible, because each phoneme in the linguistic level can be realized into numerous sounds (allophones). The role of the perceptional phonological system is to decipher the sound wave and perceive from it the linguistic level. This system, too, is very flexible: it can perceive one phoneme even if it is stimulated by different sound waves (different allophones of one

[55] The production data of Spanish and English was collected from L. Lisker and A. S. Abramson (from "Voice-timing Perception," p. 4 and from "Cross-language Study"). The production data of Hebrew was taken from Part I of Appendix 2 of this paper.

[56] E.g., J. L. Flanagan, *Speech Analysis, Synthesis and Perception* (Berlin: Springer-Verlag, 1972).

phoneme).[57] To be sure, there is a connection between the performance and the perception systems, but they are definitely two separate systems.[58] For instance, we cannot conclude what range of VOT a person who performs the phoneme /k/ in a typical context with a VOT of +57 ms is going to need in order to decipher/perceive /k/ in the same context.

Although production and perception systems are different, it is worthwhile to compare the two sets of findings. Figure 3 displays the average VOT values of stops in natural speech for 14 Hebrew speakers. Figures 6–8, which illustrate the perception of Hebrew stops, demonstrate that there is a correlation between production and perception. The fact that the positive VOT of /k/ extends longer than those of /p/ or /t/ (57 versus 27 and 24 ms) is also seen in the perception results. The dividing line from which a majority of Hebrew listeners perceive voiceless stops has a greater value for /k/ than do similar dividing lines for /p/ and /t/: most Hebrew listeners perceive /p/ or /t/ if VOT is greater than +10 ms, but they perceive /k/ only if VOT is greater than +20 ms. Abramson & Lisker found similar results in their Spanish and English perception tests,[59] although we must admit that the differences in the values between /k/ and /p t/ in each language do not seem statistically significant.

The perceptual 50% crossover points of Hebrew listeners lie in the vicinity of +14 ms, which is within the range of /p/ and /t/ but to the "left" of the range of /k/ (17–107 ms), albeit within the majority's range for the perception of /k/. A similar phenomenon occurs among English listeners: The categorical boundary of /k/ is 42 ms, which is to the "left" of the range of /k/ (50–135) but within the majority's range for the perception of /k/.

---

[57] I dealt with the performance and perception systems in more detail in A. Laufer, *Phonetics and Phonology* (Tel-Aviv: The Open University, 1992), units 4–5 of the course "Introduction to Linguistics," pp. 18–20 and 67–85 (Heb.). More on this subject can be read in numerous works; for example, D. B. Fry, *Speech Reception and Perception, New Horizons in Linguistics*, ed. J. Lyons (Harmondworth: Penguin, 1970), pp. 33–34; A. J. Fourcin, W. A. Ainsworth, G. Fant, O. Fujimura, H. Fujisaki, J. W. Hess, J. N. Holmes, F. Itakura, M. R. Schroeder, and H. W. Strube, "Speech Processing by Man and Machine," in *Recognition of Complex Acoustic Signals*, ed. T. H. Bullock, (Life Science Research Report, Vol. 5 Dahlem Workshop; University of California, San Diego, 1977), pp. 307–351; J. Lehotonen and P. Hurme, "The Speech Chain and the Science of Speech," *Papers in Speech Research*, University of Jyväskylä, Vol. 2 (1980) 1–28; M. Studdert-Kennedy, "Speech Perception," in *Contemporary Issues in Experimental Phonetics*, ed. N. J. Lass (New York: Academic Press, 1976), pp. 243–293; or G. J. Borden and K. S. Harris, *Speech Science Primer*, pp. 171–213.

[58] J. L. Flanagan, *Speech Analysis*, p. 308.

[59] A. S. Abramson and L. Lisker, "Voice-timing Perception," Figure 2; A. S. Abramson and L. Lisker, "Relative Power of Cues," Figure 3.1.

Crossover points for Spanish listeners are within the range of /p/ and /k/ but, oddly enough, to the "right" of the range of /t/.[60]

Comparing listeners' responses with the various synthetic stimuli confirms the perceptual significance of voice timing. Tests with Hebrew, English, Spanish, and Thai listeners demonstrate that this feature is perceptually relevant and sufficient.

4.0 Voicing in Contemporary Hebrew—Summary and Discussion

Our main purpose has been to check phonetic voicing in contemporary Hebrew. We did this in three stages. First, we wanted to verify that the phonological criterion that is customarily named "voicing" exists in contemporary Hebrew. Since controversies about "voicing" pertain mostly to stops, we limited our experiments to those sounds.

Secondly, we wanted to identify the acoustic features that characterize this "voice" criterion in Hebrew. Lisker and Abramson have pointed out that voice timing can serve as a very effective basis for assigning a stop to its proper linguistic category.[61] They suggest that VOT is a universal quality and have demonstrated it in eleven languages. As sounds in final position cannot be measured with the ordinary VOT method, we added a similar method called ViC (Voice into Closure).

We checked voice timing both in natural running Hebrew and in structured texts. The latter were composed so as to include an appropriate and variegated representation of Hebrew stops. At this stage, our fourteen subjects comprised the two main pronunciations of Hebrew. It is clear from our three recordings that the acoustic feature of voice timing can serve as a distinct separator for the phonemically voiced-voiceless contrast in Hebrew, with voice sounds always having a considerably longer time of vocal fold vibrations than voiceless sounds. In a non-final voiced stop, voicing starts well in advance of the release (with VOT about -90 ms) and considerably longer than before it starts in the voiceless counterpart, where VOT has a moderate lag. In a final voiced stop, voicing usually extends until the end of the sound, long after the cessation of voicing in its voiceless counterpart. (In a final voiceless stop, ViC is usually less than 10 ms, but more than a 100 ms for voiced stops.) From the voice timing perspective, there are no dialectal or individual differences in contemporary Hebrew.

---

60 A. S. Abramson and L. Lisker, "Relative Power of Cues."

61 L. Lisker and A. S. Abramson, "Cross-language Study"; "Distinctive Features."

A comparison between our findings for Hebrew and those for twelve other languages revealed similarities and differences between them. Figure 5 reveals that in all thirteen languages checked the voicing starts in the voiced sound earlier than in its voiceless counterpart; voice timing is therefore an effective measure to differentiate homorganic stop categories. We also found that Hebrew speakers resemble Spanish and Polish speakers in respect to the onset of voicing. In these three languages the span between voiced and voiceless sounds is greater than in other languages checked, and all have a long voicing lead. We also found that the moderate aspiration of Hebrew speakers resembles that among Polish speakers. We should note that a span of similar length exists also in Hebrew final stops: voicing in a final voiced stop extends much longer than in its voiceless counterpart.

Theoretically, the existence or absence of phonetic voicing does not prove that it causes the categorical differentiation of 'voicing.' An additional physical factor, such as force, air pressure, or muscle tension could serve as an acoustic cue to the voiced-voiceless contrast in Hebrew. In fact, some scholars have argued that "force" is the essential differentiating factor in Hebrew and that voicing is only an accompanying feature.[62] The third stage in our research was devoted to proving that voice timing is a distinctive and sufficient feature in Hebrew.

We proved this with a psychoacoustic test in which we synthesized three sets of Hebrew words, named "bi-pi," "dam-tam," and "gam-kam." Each set contained a series of thirty-one synthesized words that were identical in all acoustical components except for voice onset time. The timing of this voice onset was varied in 10 ms steps from -150 ms (voicing lead) to +150 (voicing lag). The order of words in each series was random. 190 Hebrew listeners identified the words in the three series. The results of this perception test revealed that Hebrew listeners have no difficulty perceiving the criterion of voicing. Usually, a sound the voicing of which started a while after the release (VOT more than 20 ms) was perceived as voiceless, while a sound with less than +10 ms of VOT was perceived as voiced. Since all other acoustical components of the "words" were identical, we were able to conclude with certainty that the cue of voice timing suffices for Hebrew speakers to differentiate phonemically between the voiced-voiceless categories. These findings are in accordance with the results of similar percep-

[62] D. Tene, "Measured Duration," §§ 4.6, 5.3; M. J. Chayen, "Pronunciation of Israeli Hebrew" pp. 288, 290; M. J. Chayen and Z. Dror, *Introduction to Hebrew*, p. 294.

tion tests of English, Spanish, and Thai listeners carried out by Abramson and Lisker.[63]

4.1 Discussion

One of the purposes of phonetic research is to discover the physiological mechanisms or the acoustic features that cause phonological differentiations. In the present research we showed a clear connection between the phonological criterion of voicing and the phonetic-physical field. Our findings in Hebrew show clearly that the physical properties of the voice timing measure (VOT and ViC) can serve as a powerful differentiator for the phonological category of voicing; indeed Figure 5 demonstrates that phonological voicing is tied to VOT, a phonetic entity, in all thirteen languages checked. Physiological voicing extends longer in sound produced as voiced than in their voiceless counterparts. Regardless of the particular language, phonemically voiced sounds always have more voicing (laryngeal activity) than their voiceless counterparts.

Languages can differ from one another (beside the number of categories along the voice timing scale) in what we called the width of the "range" of "Voice Timing," by which we mean the time range between the two left dots of each language in Figure 5. These differences seem to be quantitative, not qualitative. For example, the width of the range between voiced and voiceless sounds in Hebrew is much larger than the averages in English or Cantonese.[64]

Voice timing is, in fact, the timing of laryngeal activity called "voicing." With Lisker and Abramson, Abramson, Catford, and Flege,[65] we hold that this general term underlies a complex set of interrelated, intersecting, and overlapping acoustic and physiological features, any one of which may have perceptual efficacy. These features include the onset and offset of voice pulsing, perturbations of fundamental frequency, intensity of plosive release, amount of aspiration noise, attenuation of the first formant, duration of plosive, volume of vocal tract, and air pressure and ex-

---

63 A. S. Abramson and L. Lisker, "Voice-timing Perception"; "Relative Power of Cues."

64 We argued (A. Laufer, "'Voiced / voiceless'") that the motive for searching a new criterion to replace "voicing" was the fact that in the initial stops /b d g/ physical vibration of the vocal folds was occasionally missing. Had the researchers examined Hebrew instead of English, this problem would not have arisen. Another difference between languages is the placement of the "range" on the 'voice timing' scale—close to the release of the sound or far from it.

65 L. Lisker and A. S. Abramson, "Distinctive Features"; A. S. Abramson, "Laryngeal Control"; J. C. Catford, *Fundamental Problems*; J. E. Flege, "Laryngeal Timing and Phonation Onset in Utterance-initial English Stops," *Journal of Phonetics* 10 (1982) 177–192.

pansion of the pharynx. These features intersect in various combinations to furnish the phonetic basis for what we call phonological voicing. Therefore, speaking of voice timing, we mean all or part of these descendent features.

Although we measured voice timing by measuring the phonetic entity of the voice bar, Hebrew speakers need not be able to detect the voice bar during normal closure of a plosive. It is possible that all or part of the pulses originating at the glottis are below the threshold of audibility. (This is one reason why we did not pay attention to the intensity of the voice bar and why we treated this part of the sound as "voiced," even though we couldn't find any trace of glottal pulsation [see 2.1].)

The tests described here do not permit us to certify which is the physical clue (or the combination of clues) that causes Hebrew listeners to perceive "voice." We may assume that they use the duration of aspiration clue for the voiced-voiceless contrast. We have seen that voiced plosives are usually perceived if VOT is less than +10 ms (Figures 6–8); therefore, it is reasonable to assume that short aspiration is the clue for a voiceless sound and lack of aspiration (or VOT of less than +10 ms) for a voiced sound. Perhaps Hebrew listeners use only these clues for the voiced-voiceless contrast. But we use the general phonetic term of "voice timing" because aspiration is not a separate factor from voice timing, which is a direct consequence of the laryngeal timing,[66] and because we cannot certify which are the physical clues that cause the perception of "voice."

The ability to perceive phonological "voicing" in languages depends on the relativity of the physiological and physical entities of "voice timing." In contrast to "force," which could not be supported physiologically or physically,[67] "voice timing" was found to be an actual physiological feature of the vocal folds, and it has real acoustic effects. From our findings as well as those from twelve other languages cited here, we can suppose that the phonetic feature of voice timing, which is still a general term, is the cause of the categorical differentiation of voicing.

As perception tests of synthetic speech in English, Spanish, Thai, and Hebrew were carried out with only one aspect of voice timing changed (duration between release and voice onset), at least for these languages we can be certain that this aspect of voice timing can serve as a sole cause for

---

66 E.g., J. C. Catford, *Fundamental Problems*, p. 115.

67 E.g., L. Lisker and A. S. Abramson, "Cross-language Study"; "Distinctive Features"; and lately in A. Laufer, "'Voiced / voiceless.'"

the categorical differentiation of voicing. Although perception tests of voice timing were not actually made (as far as we know) in other languages, we may expect that this general feature of voice timing is the main differentiator of the voice-voiceless contrasts. After all, there is no evidence for the existence of physical features like "force," "muscle tension," "precision," or "lung air pressure," nor proof that the criterion of voice timing is not present in these languages. In light of the findings we mentioned and the lack of other physical features, we have no reason to hypothesize the existence of features not proven to exist. It is therefore reasonable to assume that the voice timing feature exists as a differentiating criterion[68] in all languages.

---

[68] Let us not forget that phonetic voicing (= phonation, vocal fold vibration) is the main physical sound source during speech (e.g., J. M. Pickett, *The Sounds of Speech*, p. 57, or A. Laufer, *Acoustic*, pp. 16–19). This study reveals that the timing of this voicing is also the main physical differentiator for the phonemic voice category.

APPENDICES

Appendix 1

I. Word list in the first perception test

A. When the phoneme is in an initial position, in a stressed syllable:

/k/	/g/	/t/	/d/	/p/	/b/
/ˈkever/	/ˈgever/	/ˈtofi/	/ˈdofi/	/ˈpoker/	/ˈboker/
'grave' קבר	'man' גבר	'toffee' טופי	'flaw' דופי	'poker' פוקר	'morning' בוקר
/ˈkeʃer/	/ˈgeʃer/	/ˈtoʔar/	/ˈdoʔar/	/ˈpetsa/	/ˈbetsa/
'connection' קשר	'bridge' גשר	'degree' תואר	'mail' דואר	'wound' פצע	'greed' בצע

B. When the phoneme is in an initial position, in an unstressed syllable:

/k/	/g/	/t/	/d/	/p/	/b/
/ka'baj/	/ga'baj/	/ti'ra/	/di'ra/	/po'ne/	/bo'ne/
'fireman' כבאי	'collector' גבאי	'castle' טירה	'flat' דירה	'turns' פונה	'builds' בונה
/ka'ron/	/ga'ron/	/ta'la/	/da'la/	/po'let/	/bo'let/
'cart' קרון	'throat' גרון	'hung' תלה	'drew' דלה	'emits' פולט	'stands out' בולט

C. When the phoneme is in a final position, in a stressed syllable:

/k/	/g/	/t/	/d/	/p/	/b/
/maf'lik/	/maf'lig/	/ma'rat/	/ma'rad/	/ˈʤip/	/ˈʤib/
'smacks' מפליק	'sails' מפליג	'plucked' מרט	'rebelled' מרד	'jeep' ג׳יפ	'give' ג׳יב
----	----	/ja'rit/	/ja'rid/	/ˈʤop/	/ˈʤob/
----	----	'you shot (f)' ירית	'fair' יריד	'mug' ג׳ופ	'job' ג׳וב

D. When the phoneme is in a final position, in an unstressed syllable:

/k/	/g/	/t/	/d/	/p/	/b/
/'lahak/	/'lahag/	/'zeret/	/'zered/	----	----
'squadron' להק	'idle talk' להג	'pinky finger' זרת	'twig' זרד	----	----
/'perek/	/'pereg/	/'ʃelet/	/'ʃeled/	----	----
'chapter' פרק	'poppy seed' פרג	'sign' שלט	'skeleton' שלד	----	----

<u>II. Word list in the second perception test</u>

A. When the phoneme is in an initial position, in a stressed syllable:

/k/	/g/	/t/	/d/	/p/	/b/
/'kir/	/'gir/	/'tin/	/'din/	/'pi/	/'bi/
'wall' קיר	'chalk' גיר	'clay' טין	'trial' דין	'my mouth' פי	'in me' בי
/'kam/	/'gam/	/'tam/	/'dam/	/'par/	/'bar/
'got up' קם	'also' גם	'finished' תם	'blood' דם	'bull' פר	'wild' בר
/'kuʃ/	/'guʃ/	/'tofi/	/'dofi/	/'poker/	/'boker/
'Kush' כוש	'lump' גוש	'toffee' טופי	'flaw' דופי	'poker' פוקר	'morning' בוקר

B. When the phoneme is in a final position, in a stressed syllable:

/k/	/g/	/t/	/d/	/p/	/b/
/mav'rik/	/mav'rig/	/'ʔet/	/'ʔed/	/'ʤip/	/'ʤib/
'shiny' מבריק	'screws' מבריג	'shovel' את	'steam' אד	'jeep' ג'יפ	'give' ג'יב
/'dak/	/'dag/	/'bat/	/'bad/	/ra'sap/	/ra'sab/
'thin' דק	'fish' דג	'daughter' בת	'fabric' בד	'company-major' רס"פ	'senior-major' רס"ב
/'duk/	/'dug/	/'ʃot/	/'ʃod/	/'ʤop/	/'ʤob/
'Jackstraws' דוק	'fish!' דוג	'whip' שוט	'robbery' שוד	'mug' ג'ופ	'job' ג'וב

Appendix 2

Part I sums up the VOT (in initial or in medial positions), and Part II sums up the ViC of final stops in Hebrew words (in absolute-final or in a final positions in the word, but in medial positions in the phrase).

I. VOT in ms of initial or medial stops in contemporary Hebrew
(data collected from the pronunciation of 14 informants)

Phoneme	/b/	/p/	/d/	/t/	/g/	/k/
tokens						

Sum of VOT's of voiced and voiceless stops in 572 utterances

Phoneme	/b d g/	/p t k/
Average	-93	33
Range	(-213) – (-24)	(0) – (107)
No. of tokens	275	297

II. ViC in ms of final or medial stops in contemporary Hebrew
(data collected from the pronunciation of 14 informants)

Phoneme	/b/	/p/	/d/	/t/	/g/	/k/
Average	122	6	124	8	124	7
Standard Deviation	53	14	50	13	52	13
Range	(45) – (241)	(-18) – (41)	(36) – (257)	(-27) – (36)	(44) – (254)	(-34) – (38)
No. of tokens	37	37	68	86	57	61

Sum of ViC of voiced and voiceless stops in 346 utterances

Phoneme	/b d g/	/p t k/
Average	124	7
Range	(36) – (257)	(-34) – (41)
No. of tokens	162	184