



**HAL**  
open science

# Rhythm Is a Marker of Ethnicity in Modern Hebrew: Evidence from a Perception Study and Actors' Ethnicized Portrayals

Si Berrebi, Sharon Peperkamp

► **To cite this version:**

Si Berrebi, Sharon Peperkamp. Rhythm Is a Marker of Ethnicity in Modern Hebrew: Evidence from a Perception Study and Actors' Ethnicized Portrayals. *Language and Speech*, In press, 10.1177/00238309241243025 . hal-04676625

**HAL Id: hal-04676625**

**<https://hal.science/hal-04676625v1>**

Submitted on 23 Aug 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Rhythm is a marker of ethnicity in Modern Hebrew: evidence from a perception study and actors' ethnicized portrayals

Si Berrebi<sup>1</sup> and Sharon Peperkamp<sup>2</sup>

<sup>1</sup>Department of Linguistics, Tel Aviv University

<sup>2</sup>Laboratoire de Sciences Cognitives et Psycholinguistique (ENS-PSL, EHESS, CNRS)

This is a preprint of a *Language & Speech* paper. The final version is available at <https://doi.org/10.1177/00238309241243025>.

Corresponding author: Si Berrebi [siberreb@mail.tau.ac.il](mailto:siberreb@mail.tau.ac.il)

Competing interests: The authors declare none.

# Rhythm is a marker of ethnicity in Modern Hebrew: evidence from a perception study and actors' ethnicized portrayals\*

## Abstract

In Modern Hebrew, only three segmental markers are typically acknowledged as ethnically conditioned, and usage of these markers has significantly decreased in 2<sup>nd</sup> and 3<sup>rd</sup> generation speakers. Yet the sociolinguistic situation of diverging language backgrounds of 1<sup>st</sup> generation speakers, compounded with ethnic segregation in housing and the work force, seems like a fertile ground for social identification from speech. We report two studies on prosodic variation in Modern Hebrew: a perception study and a “matched-pairs” corpus study. The results of the first illustrate that even in the absence of the known segmental markers, ethnicity perception of young native speakers may still diverge between two major ethnic identities, Mizrahi (Middle Eastern) and Ashkenazi (European). The main acoustic correlate was **rhythm**, measured as the proportional duration of vowels in the utterance. In the second study, actors' speech rhythm was found to be modulated by their portrayed ethnic identity in the same direction, suggesting that this variable is socially salient – and for some speakers, controllable – enough to be involved in style shifting. This study joins a growing body of work illustrating that relatively mild rhythmic variation can contribute to social identification, and in the current case, also for ethnicity portrayal.

**Keywords:** Rhythm; Modern Hebrew; Sociophonetics; Language attitudes; Social identification; Ethnicity

---

\* We would like to thank Outi Bat El, Noa Bassel, Noam Faust, Roey Gafter, Yehudit Henshke, Shiri Lev-Ari, Erez Levon, Almog Simchon, associate editor Molly Babel and two anonymous reviewers for their helpful comments on a previous version; Elad Zlotnik and Yoav Bar-Anan for their support using Minno, Evyatar Cohen for adjusting EasyAlign to Macintosh, Inbal Abarbanel for help with data collection, and Inbar Strull for PRAAT segmentation in the acoustic analyses presented in Section 3. This research was funded by the ANR (grant ANR-17-CE28-0007-01). ACC-BY public copyright license has been applied by the authors to the present document and will be applied to all subsequent versions up to the Author Accepted Manuscript arising from this submission, in accordance with the grant's open access conditions. Supplementary materials are available at:

[https://osf.io/8pw3k/?view\\_only=762dd378e8984579b8048aa8852b32a3](https://osf.io/8pw3k/?view_only=762dd378e8984579b8048aa8852b32a3)

# 1 Introduction

While explicit folk discourse and much of the sociolinguistics literature usually cite lexical, syntactic and segmental features as sites of meaningful social variation, research on social identification from speech – and particularly when ethnicity and race are concerned – suggests that prosody is a salient cue of a speaker’s background (American English: Purnell et al. 1999, Holliday 2016, 2021; Italian vs. Spanish: de Mareüil & Vieru-Dimulescu 2006; Swiss German varieties: Leemann & Siebenhaar 2008; Danish: Hansen & Pharao 2010; see also Armstrong et al. 2022 for a special issue on sociolectal and dialectal variation in prosody). Listeners are particularly sensitive to *rhythmic* variation, as illustrated in the context of region, ethnicity or their combination (Middle Eastern vs. North African varieties of Arabic: Barkat et al. 1999, Hamdi-Sultan et al. 2004; British vs. Singapore English: Low et al. 2000; American vs. Australian English: Vicenik & Sundara 2013; British English varieties: Torgersen & Szakay 2012; American English varieties: Purnell et al. 1999, Cogshall 2008; French varieties: Fagyal 2003, 2005; Swedish varieties: Young 2021).<sup>1</sup>

The current study is an exploration of ethno-linguistic prosodic variation in Modern Hebrew. Its primary goal was to test whether, despite an overall reduction in phonological variation among young native speakers, ethnic identification from speech may still occur through prosodic features.

Since at least Blanc (1968), Modern Hebrew is typically analyzed as having two major ethnolects, along the semi-dichotomous division of Jewish ethnicity into *Mizrahi* (Middle-Eastern and North African) and *Ashkenazi* (European). Ethnicity significantly overlaps with liturgical tradition and native languages, hence its original relationship with speech production. Starting earlier than the foundation of Israel, Zionist public policies in Palestine were oriented towards ethnic segregation in the work force and in immigrant absorption (Bernstein & Swirski 1982; Swirsky 1989; Khattab 2005; Tzfadia & Yiftachel 2008; Bashkin 2017). This vocational and geographic segregation remains correlated with socio-economic status (Ducker 2006, Cohen & Leon 2008): descendants of Jewish immigrants of European origin on average still get more education and obtain higher earnings compared with their counterparts from Middle Eastern and North African origin (Lewin-Epstein & Cohen 2019; Dahan et al. 2003), and they are considerably over-represented in the local academic and economic elite (Sasson-Levy 2013 and references therein). This state of inequality has consequences for stereotypes about Mizrahis and Ashkenazis, addressed in Section 4.

Three segmental markers are associated with the Mizrahi ethnolect: the pharyngeals [ħ] and [ʕ] and the alveolar tap/trill [r], all features of most dialects of Arabic, which was the most common L1 of 1<sup>st</sup> generation Mizrahis. Usage of these markers has been on decline since at least the 1960s; today, the

---

<sup>1</sup> Rhythm is notoriously difficult to attribute to one stable correlate. The studies cited above used a variety of measures, discussed in Section 2.1.

overwhelming majority of young native speakers use the Ashkenazi inventory. That is, they produce [χ] for both /h/ and /ħ/ (Blanc 1968; Davis 1984; though see Bentolila 1984, 2002); rarely or never produce the voiced pharyngeal fricative [ʕ], defaulting to null instead;<sup>2</sup> and use the uvular voiced approximant/fricative [ʁ] instead of [r] (Yaeger-Dror 1988). These markers have been the main focus of phonetic and phonological research on language and ethnicity in Hebrew (see Gafter 2019a for a review of sociophonetic research on Hebrew).

Segmental markers of Ashkenazi speech received less attention in the literature. That is because most features of originally Ashkenazi speech have been adopted by succeeding generations and are consequently unmarked. One possible exception is a diphthong production of the (allegedly) historically long /e:/ vowel, sometimes called “*Tzere*”/ “*Tzejrej*” after the diacritic that represents it in the writing system. Although there is evidence suggesting that production of [ej] as a marginal phoneme is more prevalent in Ashkenazi speech (e.g., Matras & Schiff 2005), this category was produced as [e] invariably in our materials, and we will therefore not elaborate further on its likely contribution to ethnicity perception.

Prosodic variation received attention in two studies. In the first, describing variation in Modern Hebrew from a broad perspective, Bolozky (2002) proposes that Yemeni and Ethiopian Jews (the first ethnic group falls under the Mizrahi umbrella term, the second is usually perceived as its own ethnic category) have *pre-tonic lengthening*, making their production more “syllable-timed” compared with Ashkenazi speech. In the second, exploring the politics of identity in Israel through language, Lefkowitz (2004) proposes that the three-step tune \*HLH% is an originally Mizrahi contour: in a sample of a free conversation between two friends, the Mizrahi among the two used \*HLH% regularly and without contextual restrictions, while her Ashkenazi friend used it less and in specific contexts, for the most part when emphasizing her involvement or agency. This, according to Lefkowitz, indicates that Ashkenazi speakers tend to perceive Mizrahis as more emotionally involved and agentive, and the borrowed contour is therefore a second-order index (Silverstein 2003; Eckert 2008) of an attitude associated with Mizrahis. If this is indeed the case, in an “out of the blue” scenario it is more likely that the contour would be used by a Mizrahi speaker.

Given the contribution of prosody to social identification in other languages, we hypothesized that if prosodic variation is indeed modulated by ethnic identity, it will likely draw listeners’ attention. Study 1 was designed to test this hypothesis.

---

<sup>2</sup> The same reduction to null happens with most occurrences of the glottal stop and glottal fricative [h], though this point of variation hasn’t been associated with ethnic identity in the literature.

## 2 Study 1: Ethnicity identification in ~5 seconds

Study 1 was inspired by Purnell et al.'s (1999) work on ethnic identification from speech in American English. Their goal was to test whether listeners can identify accents associated with White, African American, and Hispanic speakers on the phone, to test the claim that linguistic profiling is practiced in housing discrimination in the US, preventing potential residents from even viewing a property. Their results showed that both landlords (Experiment 1) and university students (Experiment 2) discriminated well between dialects based on a short phrase – “Hello, I’m calling to see about the apartment you have advertised in the paper”.

Like the original work, our primary goal was to test whether listeners can reliably classify short samples of speech according to perceived ethnicity, even those that do not contain any of the segmental markers associated with Mizrahis. To answer this question, we examined to what extent listeners agree in their judgements. Given a positive answer, a secondary goal was to detect correlations of listeners’ judgments with acoustic features. We focused on two prosodic correlates, variants of the indicators of Mizrahi speech proposed by Bolozky (2002) and Lefkowitz (2004).

### 2.1 Methods

**Participants:** One hundred and ten native Hebrew speakers volunteered to participate online. Five of them reported in the post-task comments section that they did not read the instructions well. They were excluded from the analysis, leaving 105 participants. Self-reported demographic data are summarized in Table 1. Participants labelled as “Mixed” indicated that their origins did not include only one of the identities. Those labelled “Other” indicated that neither of the umbrella terms “Mizrahi” and “Ashkenazi” applies to their background. This data was collected via a multiple-choice post-task question; no further data was collected regarding the “Mixed” and “Other” participants, and there are various possible options: for the “Other” group, participants could be immigrants from the former USSR or Ethiopia; “Mixed” could be of both Mizrahi and Ashkenazi backgrounds, or a combination of either with another group. Gender categories included “man”, “woman”, “other” and “prefer not to say”.

Ethnicity	Age (sd)	Women	Men	Prefer not to say
Ashkenazi	36 (9.8)	18	14	
Mizrahi	33 (8.3)	40	3	1
Mixed	34 (6.9)	19	4	
Other	37 (11.6)	5	1	

**Table 1: Self-reported demographic data of participants.**

Recruitment was done through Facebook posts in various groups. The skewed sample – in particular, an over-representation of Mizrahi women – is due in part to this dissemination method: one of the Facebook groups in which the experiment was published was a feminist Mizrahi group. It is likely that this group has contributed many of the Mizrahi and Mixed participants, though it is impossible to know for certain, since participation was anonymous. The current design does not allow us to test hypotheses regarding a possible effect of the listener’s background on their ratings.

**Stimuli:** Twenty-eight native Hebrew speakers were recorded producing the utterance in (1). Hebrew has grammatical gender marking on the verb, which was produced or not depending on the speaker’s gender; this marker is shown in parentheses. The utterance is shown twice: (1a) includes only the Ashkenazi-inventory segments, used by most speakers in our sample as well as by 2<sup>nd</sup> generation and onward speakers more generally, while (1b) introduces the segmental markers (in boldface) as produced with the Mizrahi inventory. Six speakers, all from the baseline older group, produced at least one of these segmental markers.

(1)

a. halo?    ʃalom,    ani mitkaʃer(et).    legabe a- diza    be-keʁov    a-avoda. (with Ashkenazi inventory)

b. halo?    ʃalom,    ani mitkaʃer(et)    legabe a- diza    be-**re**hov    a-**ʃ**avoda. (with Mizrahi inventory)

Hello?    Hello,    I call    (FSG) about Det-apartment    in-street    Det-avoda.

‘Hello? hello, I’m calling about the apartment in haavoda street.’

Speakers were of varied linguistic and ethnic backgrounds, divided across age and gender identity, the latter being inferred from whether they pronounced the sentence with or without the feminine gender marker. None of the speakers expressed discomfort or used strategies to avoid having to choose between masculine or feminine morphology. The target group consisted of 14 young speakers (aged 24-37), none of whom used the segmental markers. A baseline group, consisting of older speakers (aged 48-72), was expected to get the most coherent ratings, with relatively high levels of certainty and high rates of inter-listener agreements.

During the recording, the speaker was asked to memorize and then produce the sentence as naturally as possible without looking at the screen on which it was presented. Three productions of each speaker were recorded and the most natural-sounding one according to the first author, who is a native speaker of Hebrew, was selected.

**Acoustic features of the stimuli:** Speech samples were manually segmented by the first author on PRAAT (Boersma 2006), based on visual and auditory inspection. In what follows, we describe the

measures and how each was calculated.

The segmental markers [h], [ʕ] and [r] were identified by ear since their Ashkenazi counterparts are significantly different acoustically. Of the fourteen speakers in the Baseline (older) group, six produced at least one of the Mizrahi segmental markers [h], [ʕ] and [r]. Perhaps surprisingly, the only speaker who used [ʕ] was of European background (see discussion in 2.2.1). The other Mizrahi segmental markers were distributed according to expectation, with two older Mizrahi speakers producing [h] and [r], and three additional older Mizrahi speakers producing only [r].

Measurements of intonation contour and rhythm were conducted as follows. All speakers naturally produced the sentence as three distinct intonation phrases: (I) halo? (II) jalom, (III) ani mitkafek (et) legabe adika be-vechon aavoda.<sup>3</sup> Phrases (I) and (II) included two vowels and two or three consonants, respectively; some speakers did not produce *h* word initially in the word “hello”, in which case phrase (I) included only one consonant. The difference in the ranges of variation in duration between them was therefore dramatic (I: 81-90% V, II: 42-53% V). The third intonation phrase was significantly longer and was therefore taken to match the range of rhythmic variation in natural sentences more closely (III: 46-55% V). Accordingly, prosodic measurements are reported only for this phrase.

Contours were classified categorically as Falling, \*LH%, or \*HLH%, the third being used more consistently by Mizrahis according to Lefkowitz (2004). A contour was labeled as \*LH% if the boundary tone was the highest peak of the utterance, while it was labeled \*HLH% if the highest nucleus peak was right after pitch reset. Overall, most older speakers used a Falling contour, while young speakers used either of the rising contours. Mizrahi speakers used the \*HLH% contour more than Ashkenazi speakers ( $\chi^2(2, N = 28) = 1.8, p < .05$ ). While Ashkenazi speakers from the older group never used \*HLH%, some young speakers did, in line with Lefkowitz’s (2004) observation that this contour may be spreading.

Table 2 summarizes the dimensions of variation described so far in the recordings, by ethnicity and age group.

	Mizrahi background		Ashkenazi background	
Age group	24-33	53-72	24-33	53-72

<sup>3</sup> Women had one more syllable compared with men due to gender marking, which includes both a consonant and a vowel (/et/). Men and women were distributed equally across ethnic backgrounds; in addition, gender did not affect the results significantly in general and more particularly, in the relevant intonation phrase, as indicated by a t-test comparing men vs. women’s %V:  $t(26) = 1.5, p > 0.1$ .



[h]	0/7	2/7	0/7	0/7
[ʃ]	0/7	0/7	0/7	1/7
[r]	0/7	5/7	0/7	0/7
*HLH%	5/7	3/7	2/7	0/7

**Table 2: Number of speakers that exhibit possible ethnicity markers, by speaker’s ethnic background and age group. \*HLH% refers to the contour of intonation phrase III. Rhythm data is presented separately in Figure 3.**

The next dimension of variation is the most challenging to measure but – as alluded to in the introduction – also the most promising: rhythm. For Hebrew, Bolozky (2002) suggested that pre-tonic lengthening may distinguish some speakers on the ethnicity axis; however, it is not clear what environment should be considered for pre-tonic lengthening – every pre-tonic syllable, only open syllables, or perhaps only syllables that bear a pitch accent? An in-depth production study is required to pin down the environment for this proposed phonological process. In the meantime, we opted for a more general approach, searching for a rhythm measure that would reflect variation in the entire intonation phrase. (We return to some of the potential underlying reasons for rhythmic variation in Modern Hebrew, including pre-tonic lengthening, in Section 4).

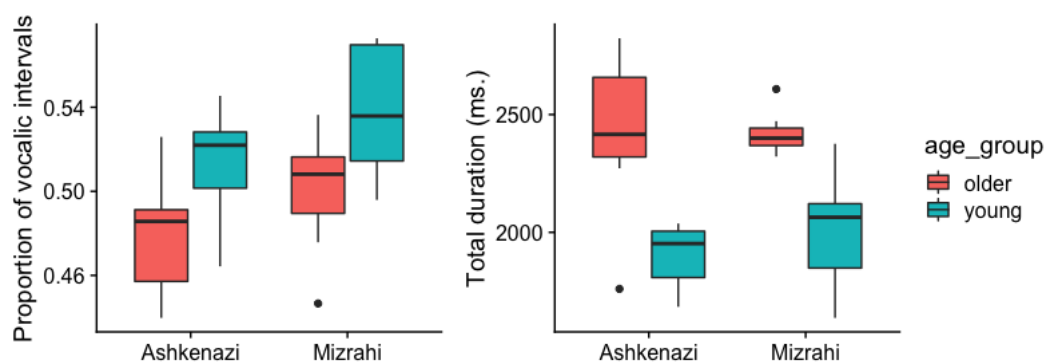
Listeners are sensitive to rhythmic features of speech. For example, Ramus & Mehler (1999) demonstrated that French listeners can rely on rhythmic properties alone (i.e., when segmental and pitch cues are made unavailable through speech synthesis) to achieve above-chance discrimination between English and Japanese. At the same time, rhythmic perception is notoriously difficult to tie consistently with the speech signal: using a corpus of read text from multiple speakers in five different languages (Russian, Mandarin, English, Greek, French), Loukina et al. (2011) found differences in performance of 15 different rhythm measures, with none being broadly successful in language discrimination.

The various measures for rhythm proposed in the literature differ in whether they are calculated over vocalic intervals (Grabe & Low 2002), feet (Nolan & Asu 2009), or a combination of vocalic intervals and variability in consonantal intervals (Ramus et al. 1999). Rhythm perception is likely to be affected not only by raw (or relative) duration, but also by pitch (Barry et al. 2009) and by the degree of vowel reduction (Low et al. 2000). We used the total duration of vocalic intervals, divided by the total duration of all segments in the intonation phrase (hereafter %V). We chose this measure based on evidence that it is strongly correlated with language discrimination: modelling the results of the English-Japanese rhythmic discrimination task described above, Ramus et al. (1999) found that the lower the %V value of an English sentence, the better it was classified by subjects (that is, sentences with higher %V were more confusable

with Japanese sentences). Perceptual results from Arabic dialect identification point in the same direction: given synthesized stimuli that included only prosodic cues, discrimination between Eastern and Western dialects was above chance (Barkat et al. 1999); the dialect groups are distinct on the %V axis (Hamdi Sultan et al. 2004).

There were no glides in the recorded sentence, so classification into consonants and vowels was straightforward. Despite the relatively small sample size, there was a trend such that %V was higher for Mizrahis ( $t(25.82) = -1.85, p = 0.076$ ).

The rhythm measure %V was correlated with other measures: first, with usage of the \*HLH% contour, such that speakers who used the \*HLH% contour also tended to have a higher %V (Pearson's correlation:  $R(26) = 0.46, p = 0.01$ ); second, with the total duration of the phrase ( $R(26) = -0.50, p = 0.007$ ). The main predictor of total duration was age, such that older speakers talked more slowly than younger speakers, regardless of ethnic background ( $t(24.82) = -5.57, p < .001$ ). This is consistent with previous acoustic studies on American English read speech, in which younger adults were found to speak faster than older adults (Smith et al. 1987, Amerman & Parnell 1992). Total duration did not distinguish between ethnic backgrounds on its own ( $t(24.22) = -0.41, p = 0.68$ ). Figure 3 shows the %V and the total duration as a function of ethnicity and age group.



**Figure 3: %V and total duration of intonation phrase III, by ethnic background and age group.**

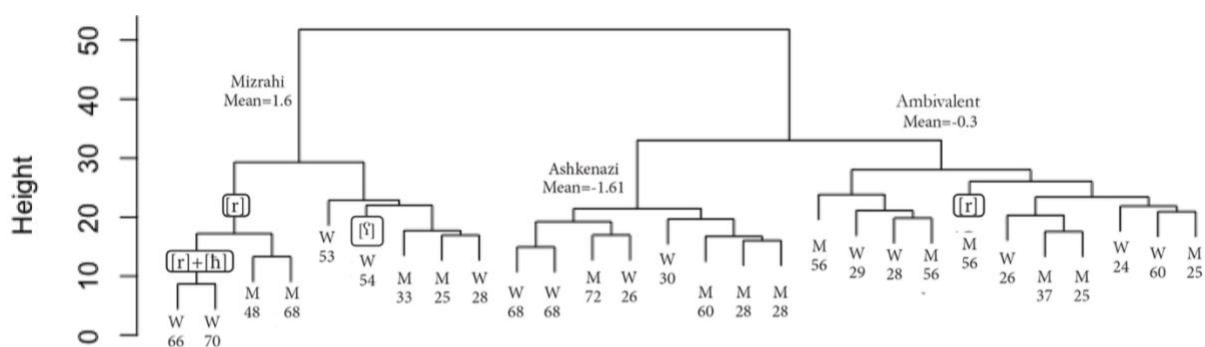
**Procedure:** The study ran on the online platform *Minno* (Zlotnick et al. 2015). Before the task, participants filled in a form collecting demographic data: age, gender, native language(s), and main place of residence since childhood. In the main task, “Mizrahi” and “Ashkenazi” were presented at opposite ends of a 7-point scale, such that the mid-point was “I can’t tell” and. Participants were requested to wear headphones, listen to each recording, and use the scale to mark their level of certainty. Each recording appeared on a separate screen, and participants could listen for as many times as they

wished by pressing a button, before moving on to the next screen. The recordings were presented in random order. After the judgment task, participants completed a short questionnaire with more demographic and ideological questions. For a complete presentation of the exit questionnaire, see Berrebi (2021).

## 2.2 Results and discussion

### 2.2.1 Inter-listener agreement

Ratings were coded from -3 (more Ashkenazi) to 3 (more Mizrahi), with 0 corresponding to the response “I can’t tell”. Hierarchical clustering was applied to the perceived ethnic identity ratings of both the target (young) and baseline (older) speakers, using the `hclust` function in R (R Core Team 2017). This method creates additive similarity trees, in the current case assessing the perceptual distance between talkers’ perceived ethnicity. Here, the rationale is that voices who use the segmental markers will be clearly rated as Mizrahi; if young speakers cluster together with these voices despite the absence of the segmental markers, we can safely conclude that they are perceived as Mizrahi as well. Initially, each recording is assigned a separate cluster; then the algorithm proceeds iteratively, at each stage joining the two most similar clusters, until there is just one. At each stage, distances between clusters are recomputed by the Lance–Williams dissimilarity update formula, using the complete linkage method. In terms of presentation, each resulting subtree is ordered such that the tighter cluster in each bifurcation is on the left.



**Figure 4:** Hierarchical clustering of perceived ethnicity, based on 105 ratings per recording. Each leaf represents a speaker, with age and gender noted. The branches containing known Mizrahi markers are labeled. Height represents Euclidean distance.

In the dendrogram presented in Figure 4, the best choice for the total number of clusters by eye is either two or three. For convenience, the clusters can be labelled “Mizrahi”, “Ashkenazi”, and

“Ambivalent” (in the case of three clusters), or “Mizrahi” and “Ashkenazi-ambivalent” (in the case of two clusters). To estimate internal consistency and decide whether the better cut is into two or three clusters, Cronbach’s alpha was computed for each. For the Mizrahi cluster (9 recordings), reliability was acceptable to good ( $\alpha = 0.7, CI = 0.62 - 0.79$ ). For the Ashkenazi cluster (8 recordings), reliability was questionable to good ( $\alpha = 0.63, CI = 0.53 - 0.74$ ), and for the ambivalent cluster (11 recordings) the results were unreliable ( $\alpha = 0.37, CI = 0.2 - 0.55$ ). With a two-cluster cut, reliability for the Ashkenazi-ambivalent cluster was worse than for the Ashkenazi cluster ( $\alpha = 0.59, CI = 0.47 - 0.7$ ).

Table 5 includes the mean and standard deviation for all voices, grouped into the clusters found by the algorithm. Importantly, the resulting Mizrahi cluster included voices from the baseline older speakers group, but also speakers from the target young group who do not produce the known segmental markers. This confirms that there are significant differences in ratings of recordings that cannot be based on the known, segmental, markers.

Still, the results also confirm that the segmental markers, and especially [h], are the least contested aspect of ethnically-conditioned variation. Usage of [h] reduced variation between listeners’ judgements (see standard deviation for speakers who used this marker in Table 5). The only speaker in the present study who produced [ʕ] received mixed responses, as evident by the relatively high standard deviation. These findings are in line with Gafter (2016b), who found that [h] is used systematically by speakers who maintain the distinction between [χ] and [h], contrary to [ʕ], which is sensitive to style and register and may be used by Ashkenazi speakers in reading contexts. Similarly, [r] is used by speakers of various languages in addition to native speakers, including Arabic (closely related to the Mizrahi identity) but also European languages (associated with an Ashkenazi identity). There were three speakers who produced [r] but no other segmental markers; of them, two were rated as Mizrahi relatively reliably. The third, who is a heritage speaker of Iraqi Arabic, had the highest overall SD and was clustered by the algorithm with the “ambivalent” voices. There may have been other features of the recording that sounded more Ashkenazi, explaining why listeners tended to use both sides of the scale (contrary to hovering around the mid-point of the scale, which was the case for most “ambivalent” voices).

Cluster	Age	Gender	Mean	SD	Markers
Mizrahi	70	W	2.63	0.82	[h] and [r]
Mizrahi	66	W	2.5	1.14	[h] and [r]

Mizrahi	68	M	2.04	1.4	[r]
Mizrahi	48	M	1.67	1.27	[r]
Mizrahi	25	M	1.58	1.32	
Mizrahi	28	W	1.23	1.34	
Mizrahi	33	M	1.21	1.36	
Mizrahi	54	W	1.05	1.8	[ʃ]
Mizrahi	53	W	0.48	1.54	
Ashkenazi	68	W	-2.24	0.95	
Ashkenazi	68	W	-1.98	1.37	
Ashkenazi	28	M	-1.69	1.33	
Ashkenazi	60	M	-1.65	1.35	
Ashkenazi	28	W	-1.66	1.4	
Ashkenazi	72	M	-1.52	1.47	
Ashkenazi	28	M	-1.33	1.31	
Ashkenazi	30	W	-0.84	1.39	
Ambivalent	60	W	-0.74	1.6	
Ambivalent	26	M	-0.65	1.5	
Ambivalent	24	W	-0.55	1.6	
Ambivalent	56	M	-0.50	1.9	[r]
Ambivalent	26	W	-0.44	1.6	
Ambivalent	56	M	-0.28	1.4	
Ambivalent	29	W	-0.22	1.6	
Ambivalent	37	M	-0.19	1.4	
Ambivalent	28	W	-0.12	1.5	
Ambivalent	24	M	0.21	1.3	
Ambivalent	56	M	0.17	1.8	

**Table 5: Individual speaker’s age, gender, mean ethnicity rating, SD of ethnicity rating, and usage of segmental markers, arranged by cluster. The ethnicity ratings are on a scale from -3 (more Ashkenazi) to 3 (more Mizrahi).**

To reiterate our main point, while none of the 14 young speakers used any of the known segmental markers associated with ethnicity, half of them were still identified as Mizrahi (N=3) or Ashkenazi (N=4), based on the clustering analysis. The Mizrahi-identified ones were all from Mizrahi-majority neighborhoods; among the Ashkenazi-identified ones, three were from Ashkenazi majority neighborhoods, and one was from a Mizrahi-majority settlement (see supplementary materials for full details).

### 2.2.2 Acoustic correlates of the ratings

Variation in the production of known segmental markers only occurred in the older group. That is, the potential contribution of most of the predictors for older speakers' ratings to the rating of younger speakers is null. Therefore, linear mixed-effects regression models were fit separately by age group to assess the contribution of acoustic properties to the dependent measure, i.e., the ethnicity ratings. Each model was initially maximally specified, and fixed effects that did not significantly contribute to data likelihood (measured as the Bayesian Information Criterion) were gradually removed.

The model for the baseline group of older speakers initially included the fixed factors [r] and [ʕ] (ʁ/r and ʕ/null; both dummy-coded), Contour (\*HLH% / Other, dummy-coded), and Rhythm (continuous). Voice (speaker) and Participant were included as random intercepts. Due to collinearity with [r], [h] was not included in the model (all speakers who produced [h] also produced [r], but not the other way around). Contour did not contribute to the model and was therefore removed. Replacing the rhythm measure with total duration worsened the model. The final model is summarized in Table 6.

The older speakers model confirms the observation from previous studies that the tap/trill [r] variant and the pharyngeal [ʕ] contribute to ethnicity perception from speech. It adds the insight that rhythm may play a role in ethnicity perception as well. It should be noted that while [h] was removed from the model due to collinearity, previous studies on Modern Hebrew suggest that it substantially contributes to ethnicity perception (Blanc 1968, Davis 1984, Bentolila 1984, Gafter 2016a, 2019b); in the current study, the two speakers who produced [h] in addition to [r] had the most extreme Mizrahi ratings.

The model for the target group of young speakers included Rhythm and Contour as fixed effects, in addition to random intercepts for Voice and Participant. Contour did not contribute to the model fit and was therefore removed. The final model is summarized in Table 7.

---

#### Random effects

---

Groups	Name	Variance	Std. Dev.
--------	------	----------	-----------

---

Participant	(Intercept)	0.084	0.29		
Voice	(Intercept)	0.39	0.62		
Number of observations: 1470, Number of groups: Participant, 105; Voice, 14					
<b>Fixed effects</b>					
	$\beta$	<i>SE</i>	<i>Df</i>	<i>t</i>	<i>p</i>
(Intercept)	-13.96	2.92	14	-4.78	< 0.001
r	2.14	0.38	14	5.61	< 0.0002
ç	2.38	0.68	14	3.5	< 0.01
Rhythm	27.11	6.08	14	4.46	< 0.001

**Table 6: Model output for optimal model predicting ethnicity ratings for Baseline group (older speakers).**

<b>Random effects</b>	Name	Variance	Std. Dev.		
Participant	(Intercept)	0.12	0.34		
Voice	(Intercept)	0.61	0.78		
Number of observations: 1470, Number of groups: Participant, 105; Voice, 14					
<b>Fixed effects</b>					
	$\beta$	<i>SE</i>	<i>Df</i>	<i>t</i>	<i>p</i>
(Intercept)	-10.57	3.65	14	-2.8	<.05
Rhythm	19.64	6.92	14	2.84	<.05

**Table 7: Model output for optimal model predicting ethnicity ratings for target group (young speakers).**

Here too, Rhythm had a significant effect on ratings, such that a higher proportion of vocalic intervals in the third intonation phrase correlated with perceived Mizrahiness. As a sanity check, we also tested a model where Rhythm was replaced with the total duration of the third intonation phrase. Duration did not correlate with ethnicity perception.

While Contour did not contribute significantly to either model, we refrain from concluding that it does not affect ethnicity perception in Modern Hebrew. As noted by Podesva (2011), the salience of a prosodic pattern as bearing some social meaning can be based on its frequency of usage as well as on its phonetic realization. Both can be at play in the current case. First, in Lefkowitz (2004), the difference between Mizrahi vs. Ashkenazi speech was not the presence vs. absence of the \*HLH% contour; rather,

the two groups differed in their usage of this contour – regularly every several sentences vs. under specific pragmatic circumstances. Second, we did not measure F0 continuously, but assigned contour as a categorical variable. In Podesva’s (2011) English production data, the phonetic manifestation of a contour was correlated with the strength of its social meaning (that is, more extreme phonetic values were associated with clearer social meaning), which is of course missed in categorical encoding. Finally, we note that despite its non-significant contribution to ethnicity perception, the Mizrahi speakers in our sample did tend to use \*HLH% more often than the Ashkenazi speakers (see Section 2.1).

Study 1 was a first quantitative exploration of ethnicity perception from speech in Modern Hebrew, beyond the known segmental markers. It did not test one manipulated variable, but a variety of features as they are found in naturally-produced read speech. A resulting limitation is that there can be any number of acoustic variables that were not included in the analysis but were salient to listeners (e.g., voice quality), some of which might be correlated with rhythm. In other words, the study does not allow us to deduce a direct relationship between rhythm and ethnicity perception. Study 2 was designed to go beyond correlation.

### **3 Study 2: Style shifting in “ethnicity portrayals”**

A signature trait of linguistic markers is their involvement in style shifting (Labov 1973): if speakers change the rate of variant usage according to social circumstances, they must at some level be aware of the variant’s social meaning. Study 2 set out to test whether rhythm participates in style shifting, based on the correlation between rhythm and ethnicity perception found in Study 1. More particularly, we analyzed speech from characters portrayed by a single Israeli actor/actress who is a native Hebrew speaker, to examine whether rhythm varies as a function of their portrayed ethnic identity.

The use of scripted data has become increasingly common in sociolinguistic research, particularly in cases where performances are aimed at displaying a linguistic variety (Schilling-Estes 1998), embodying a stereotype (Pratt & D’Onofrio 2017; Gafter 2019b), or parodically imitating a specific person (Sclafani 2009). Performances in this context have the potential to inform the sociolinguist about markers that are perceived as important features of a group or an individual’s speech. Support for this idea comes from a perception experiment of real and imitated accented speech, in which German listeners more reliably “identified” an imitated accent produced by a native speaker, compared with a real L2 accent (Neuhauser and Simpson 2007). In the current case, performances were not necessarily parodic imitations of Mizrahi or Ashkenazi characters, although in some of the sketches – particularly those in which Mizrahi characters are portrayed – ethnicity is highlighted. If rhythm is perceptually salient, we expect that at least some of the



actors would be able to use it to display ethnicity.

### 3.1 Methods

Four actors were chosen, two women (Mizrahi: Orna Banai, Ashkenazi: Keren Mor) and two men (Mizrahi: Menashe Noy, mixed background: Shai Avivi). Of them, only Orna Banai grew up in a clearly Mizrahi-majority city (Be'er Sheva). Three of the actors – all except Orna Banai – performed together on the same sketch show from which most of the sketches used in the analysis were taken. Characters were selected for each actor according to their portrayed ethnic background. Two short segments of freely available content were used for each actor. The full audio, segmentation on PRAAT and analysis files are available in the supplementary materials.

As evident in Table 8, the display of ethnicity is supported by the character's vocation or status, i.e. typically lower for Mizrahi compared with Ashkenazi, reflecting what – as mentioned in the Introduction – is the case on average in Israeli society. The sketches diverge in the number of additional cues of portrayed ethnicity. For example, in the case of the Mizrahi character of Orna Banai, there are exaggerated syntactic and lexical stereotypes and many other cues (the character lives in a Mizrahi-majority city, she goes to visit relatives in a Mizrahi-majority town); in the sketch with the Mizrahi character of Menashe Noy, the point of the sketch is that he is Iraqi-Israeli, which is one of the identities under the Mizrahi umbrella term. On the other hand, the cues for ethnicity of the Mizrahi characters of Keren Mor and Shai Avivi, as well as most of the Ashkenazi characters, are more subtle. We assume, though, that each of the selected characters would be recognized by Israeli viewers as Mizrahi or Ashkenazi respectively. An elaborate description of the sketches can be found in the supplementary materials.

Actor	Character's vocation	Character's assumed ethnicity	Number of IPs	Mean length of IP (in segments)	%V
-------	----------------------	-------------------------------	---------------	---------------------------------	----

Orna Banai	Unknown	Mizrahi	52	24.9	48.8%
Orna Banai	Police detective	Ashkenazi	42	31.1	42.3%
Keren Mor	Supermarket cashier	Mizrahi	23	20.2	46.3%
Keren Mor	Landlord	Ashkenazi	32	28	43.6%
Menashe Noy	Unsuccessful boxer	Mizrahi	22	33.7	51.3%
Menashe Noy	Doctor	Ashkenazi	21	21.3	44.5%
Shai Avivi	Watermelon seller	Mizrahi	31	19.9	54.1%
Shai Avivi	Actor manager	Ashkenazi	27	16.5	51.1%

**Table 8: Summary of acoustic data (by actor and character/ethnicity).**

Each sketch was divided into intonation phrases by the first author, and then fully segmented on PRAAT using EasyAlign for Hebrew (Silber-Varod et al. 2022). The initial segmentation was then corrected based on visual and auditory inspection by a trained research assistant, who was not aware of the research question. Questions, hesitations, and unintelligible phrases that ended mid-sentence or involved a lot of noise and/or speech from another speaker were excluded. Vowels and consonants were separated for the analysis, with [j], the only glide in the materials, being treated as a consonant.

### 3.2 Results and discussion

The analysis included 250 intonation phrases (see breakdown according to actor and character ethnicity in Table 8). A linear mixed effects model was fitted to the data, with the fixed factors Character ethnicity (contrast-coded), Log number of segments and their interaction. The dependent variable was Rhythm (%V), as measured in Study 1. Actor was added as a random intercept. Log number of segments and their interactions were removed since they did not contribute to data likelihood.

The full model is presented in Table 9. There was a main effect of Character ethnicity in the expected direction, confirming that actors had on average a larger proportion of vowel duration when portraying a Mizrahi character. Figure 10 is a visual summary of the results. Each data point represents one intonation phrase, and the longer it is the darker its color.

Random effects	Variance	Std. Dev.
Actor (Intercept)	7.98	2.82

Residual 56.86 7.54

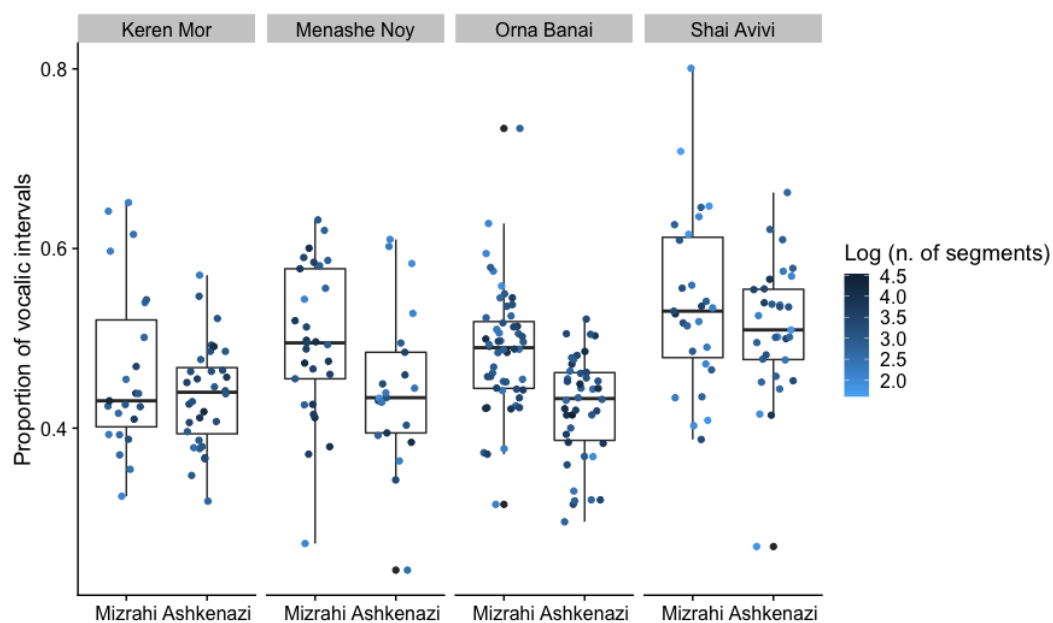
Number of objects: 250, Number of actors: 4

<b>Fixed effects</b>	Estimate	Std. Error	Df	t value	Pr(> t )
(Intercept)	47.82	1.5	4.03	31.96	<.001
Character ethnicity (= Mizrahi)	2.44	0.48	246.54	5.08	<.001

**Table 9: Results of a linear mixed effects model of rhythm as a function of portrayed ethnic identity and length of utterance.**

These results show that actors used rhythm as a marker of ethnic identity. The reasoning is that, while the viewers of a show may perceive a given character as Mizrahi or Ashkenazi based on features like clothing, posture, vocation and even syntactic or lexical variants, the actor who *portrays* the character has no reason to change the proportional duration of vocalic intervals, other than to contribute to the overall ethnic impression. This, then, suggests that the correlation of rhythm with ethnicity perception found in Study 1 is not merely a correlation: it is likely that listeners, just like actors in Study 2, used rhythm in a socially informative way.

The one actress who doesn't seem to manipulate rhythm according to the assumed portrayed ethnicity is the one from an Ashkenazi background (Keren Mor). The actress who seems to have the most reliable rhythmic distinction between her two characters, Orna Banai, was born and went to school in the Mizrahi-majority city of Be'er Sheva. Differences in the degree to which rhythm is manipulated may be due to individual differences in the actors' experience with or commitment to that aspect of the impersonation, or alternatively, to differences in the extent to which the character is meant to be portrayed as Mizrahi/Ashkenazi (e.g., Banai's Mizrahi character is made to be perceived as undoubtedly Mizrahi).



**Figure 10: Proportion of vocalic intervals given actor, ethnic identity and log number of segments in the utterance.**

## 4 General discussion

Our results illustrate that rhythm is a marker in Modern Hebrew. Several factors may contribute to the proportionally longer vowel duration in Mizrahi-perceived speech. On the one hand, Ashkenazi speakers may have higher rates of vowel reduction and/or deletion, for two reasons. First, pervasive vowel reduction in Yiddish (Faust 2018) may have had long-lasting substrata effects on Hebrew speakers from an Ashkenazi background. Second, younger Ashkenazi speakers may be leading English-aligning vowel reduction. Indeed, English proficiency is very high among some young speakers, especially those from a high SES background, who tend to be Ashkenazi; they are exposed to English (including through TV series, music and social media), and they often use English loanwords, with some also codeswitching.

On the other hand, Mizrahi speakers may have vowel lengthening in some positions (Boložky 2002). Pretonic lengthening is also characteristic of young native speakers of French in a suburb of Paris, la Courneuve, in which many young speakers are 2<sup>nd</sup> and 3<sup>rd</sup> generation to immigration from North Africa (Fagyal 2003, 2005). What seems to be contributing to the prominence of the pretonic syllable, apart from duration, is a shift of the high pitch accent – from the final syllable in the intonation phrase (in Parisian French) to the penultimate syllable. While Fagyal does not commit to the source of early pitch alignment and pre-tonic lengthening in French, she does point out that it is a feature of Moroccan Arabic (2005:101). This suggests a shared source with

Hebrew, since during the 1950s, Jewish immigrants from Morocco became the majority under the Mizrahi umbrella term (Lissak 1999; p. 9-10). We provisionally take “rhythm” as a proxy for both Ashkenazi reduction and Mizrahi lengthening and conclude that these factors should all be investigated separately in the future.

Stereotyped variants tend to occur less in formal contexts (e.g., Labov 1966); in the case of Modern Hebrew, the stereotyped status of segmental markers led to their almost complete disappearance from the speech of 2<sup>nd</sup> and 3<sup>rd</sup> generation speakers. However, as recent work – including the current study – demonstrates, ethnic identification from speech can outlive overtly stereotyped variants. In Purnell et al.'s (1999) study, even as little as a single English word – “Hello” – produced by a single speaker, allowed listeners to reliably identify the speaker's dialect as Standard American English, AAVE, or Chicano English. This has consequences at the societal level, since as King et al. (2022) recently showed, speakers can be evaluated negatively based on ethnic identification alone, even in the absence of stereotyped linguistic features. In their matched guise task, listeners judged a speaker either under a Mainstream American English (MAE) or an AAVE guise; in both, the speaker, who was a male Black professional, was mostly identified as African American. Listeners' estimations of the speaker's education attainment (college, professional degree etc.), although significantly higher in the MAE guise, were overall low in both. Further, under both guises the speaker was mostly heard as Lower or Working class, with no difference between guises. In their 1999 study, Purnell et al. found four acoustic correlates of ethnic identification, two of them prosodic; the speaker in King et al.'s (2022) study likely shared some of these features under both guises. Taken together, these findings suggest that ethnic identification from speech, based on prosody and voice quality, can be enough to induce some of the same social evaluations as the stereotyped features of an ethnolect.

In the Israeli context, our study is a first demonstration that Hebrew listeners can use a cue outside the familiar segmental markers, i.e., rhythm, for ethnic identification. The context in which actors use their “Mizrahi-leaning” and “Ashkenazi-leaning” rhythms – the former for vulgar, uneducated characters with low-paying jobs and the latter for professional positions – implies that rhythm, like the voice features of King et al.'s (2022) recorded speaker, projects particularly onto SES evaluations.

There is a tension between the sociolinguistic focus on segmental variation as the main giveaway characteristic of social belonging (or “Shibboleth”) – not only in the literature, but also and as importantly, in folk discourse – and the fact that listeners don't necessarily require overt stereotypization for social identification. This is particularly true with non-categorical features like rhythm and F0, that are typically gradient on the one hand, salient on the other. Undoubtedly, the stereotypization of specific variants plays

a major role in social identification; it is much less clear what types of linguistic awareness are involved in the utilization of non-stereotyped markers for identification. For example, aside from trained actors, do people style-shift with rhythm? As Vaughn & Kendall (2019) recently showed, when speakers are asked to produce a marked variant, they tend to also produce other variants associated with the same social group. If non-stereotyped markers are perceived as part of a “general” accent, does using (or not using) them affect perceived cohesion? Future work on these questions can elucidate the relationship between variation, awareness and control of linguistic varieties (Preston 1996; 2013).

## References

- Amerman, J. D., & Parnell, M. M. (1992). Speech timing strategies in elderly adults. *Journal of Phonetics*, 20(1), 65-76.
- Armstrong, M., Breen, M., Gooden, S., Levon, E., & Yu, K. M. (Eds.) (2022). Sociolectal and Dialectal Variation in Prosody [special issue]. *Language and Speech* Vol. 65(4), 783–790 Sage Publications UK: London, England. DOI: <https://doi.org/10.1177/002383092211221>
- Bashkin, O. (2017). *Impossible Exodus: Iraqi Jews in Israel*. Stanford University Press.
- Barkat, M., Ohala, J., & Pellegrino, F. (1999). *Prosody as a distinctive feature for the discrimination of Arabic dialects*. In Sixth European Conference on Speech Communication and Technology.
- Barry, W., Andreeva, B., & Koreman, J. (2009). Do rhythm measures reflect perceived rhythm?. *Phonetica*, 66(1-2), 78-94.
- Bentolila, Y. (1984). The sociophonology of Hebrew as spoken in a rural settlement of Moroccan Jews in the Negev. [Hebrew: Mivta’ej ha-‘ivrit ha-mešamešet be-mošav šel joc’ej maroko ba-negev]. Jerusalem: Hebrew University.
- Bentolila, Y. (2002). *Speaking Hebrew: Studies in the Spoken Language and in Linguistic Variation in Israel [Hebrew]*, chap. Linguistic variation across generations in Israel, (pp. 131–140). (Te’uda 18).
- Bernstein, D., & Swirski, S. (1982). *The rapid economic development of Israel and the emergence of the ethnic division of labour*. *British Journal of Sociology*, (pp. 64–85).
- Berberi, S. 2021. *Hearing ethnicity: classification, social stereotypization and the processing*

*of ambiguous phonetic/social cues in Modern Hebrew*. Doctoral dissertation, École Normale Supérieure and Tel Aviv University.

Blanc, H. (1968). *The Israeli Koine as an Emergent National Standard*. John Wiley & Sons, Incorporated.

Boersma, P. (2006). *Praat: doing phonetics by computer*. <http://www.praat.org/>

Bolozky, S. (2002). Phonological and morphological variation in spoken Hebrew [Hebrew: Šonut fonologit u-morfologit ba-‘ivrit ha-meduberet]. *Te’uda* 18 (pp. 239–278).

Cohen, U., & Leon, N. (2008). The new Mizrahi middle class: Ethnic mobility and class integration in Israel. *Journal of Israeli History*, 27(1), 51-64.

Coggshall, E. L. (2008). The prosodic rhythm of two varieties of Native American English. *University of Pennsylvania Working Papers in Linguistics*, 14(2), 2.

Dahan, M., Dvir, E., Mironychev, N., & Shye, S. (2003). Have the gaps in education narrowed? on factors determining eligibility for the Israeli matriculation certificate. *Israel Economic Review*, 2, 37–69.

Davis, L. M. (1984). The pharyngeals in Hebrew: Linguistic change in apparent time. *Folia Linguistica Historica*, 18 (Historica vol. 5, 1), 25–32.

De Mareüil, P. B., & Vieru-Dimulescu, B. (2006). The contribution of prosody to the perception of foreign accent. *Phonetica*, 63(4), 247-267.

Ducker, C. (2006). Jews, Arabs and Arab Jews: the politics of identity and reproduction in Israel. *ISS Working Paper Series/General Series*, 421, 1-58.

Eckert, P. (2008). Variation and the indexical field. *Journal of sociolinguistics*, 12(4), 453–476.

Fagyal, Z. (2003). La prosodie du français populaire des jeunes. *Le Français aujourd’hui*, (4), 45–55.

Fagyal, Z. (2005). Prosodic consequences of being a Beur: French in contact with immigrant languages in Paris. *University of Pennsylvania Working Papers in Linguistics*, 10(2), 8.

Faust, N. (2018). Bi-Morphemic Virtual Length in Yiddish. *Studia Linguistica*, 72(3), 608-624.

- Gafter, R. J. (2016a). Pharyngeal beauty and depharyngealized geek. *Raciolinguistics: How Language Shapes Our Ideas About Race*, (p. 185).
- Gafter, R. J. (2016b). What's a stigmatized variant doing in the word list? authenticity in reading styles and Hebrew Pharyngeals. *Journal of Sociolinguistics*, 20(1), 31–58.
- Gafter, Roey J. (2019a). Modern Hebrew sociophonetics. *Brill's Journal of Afroasiatic Languages and Linguistics*. 11: 226-242.
- Gafter, R. J. (2019b). Using pharyngeals out of context: Linguistic stereotypes in parodic performances of Mizrahi Hebrew speakers. In *The Routledge Companion to the Work of John R. Rickford*, (pp. 232–241). Routledge.
- Grabe, E., and E. L. Low. (2002). Durational variability in speech and the rhythm class hypothesis. *Papers in Laboratory Phonology 7* (515-546): 1-16.
- Hamdi Sultan, R., Barkat-Defradas, M., Ferragne, E., & Pellegrino, F. (2004, October 4–8). Speech timing and rhythmic structure in Arabic dialects: A comparison of two approaches. In *International Speech and Communication Association* (pp. 1613–1616). <https://shs.hal.science/halshs-01740967/document>
- Hansen, Gert Foget, & Nicolai Pharao (2010). Prosody in the Copenhagen multiethnolect. In Pia Quist & Bente Ailin Svendsen (eds.), *Multilingual urban Scandinavia: New linguistic practices*, 79–95. Bristol: Multilingual Matters.
- Holliday, N. R. (2016). Intonational variation, linguistic style and the Black/Biracial experience. New York University. Doctoral dissertation.
- Holliday, N. (2021). Prosody and sociolinguistic variation in American Englishes. *Annual Review of Linguistics*, 7, 55–68.
- Khattab, N. (2005). Ethnicity, class and the earning inequality in Israel, 1983-1995. *Sociological Research Online*, 10(3), 1–17.
- King, S., Vaughn, C., & Dunbar, A. (2022). Dialect on Trial: Raciolinguistic ideologies in perceptions of AAVE and MAE codeswitching. *University of Pennsylvania Working Papers in Linguistics*, 28(2), 7.
- Labov, W. (1966). *The Social Stratification of English in New York City*. Washington: Center



for Applied Linguistics.

Labov, W. (1973). *Sociolinguistic patterns* (4). University of Pennsylvania press.

Leemann, A., & Siebenhaar, B. (2008). Perception of dialectal prosody. In *Ninth Annual Conference of the International Speech Communication Association*.

Lefkowitz, D. (2004). *Words and stones: The politics of language and identity in Israel*, vol. 26. Oxford University Press.

Lewin-Epstein, N., & Cohen, Y. (2019). Ethnic origin and identity in the Jewish population of Israel. *Journal of Ethnic and Migration Studies*, 45(11), 2118–2137.

Lissak, M. (1999). The Mass Immigration in the Fifties: The Failure of the Melting Pot Policy. [Hebrew: Ha-‘Aliya ha-gdola bi-šnot ha-hamišim: kišlono šel kur ha-hitukh]. The Bialik Institute, Jerusalem.

Loukina, A., Kochanski, G., Rosner, B., Keane, E., & Shih, C. (2011). Rhythm measures and dimensions of durational variation in speech. *The Journal of the Acoustical Society of America*, 129(5), 3258-3270.

Low, E.L., Grabe, E., Nolan, F., (2000). Quantitative characteristics of speech rhythm: ‘syllable-timing’ in Singapore English. *Language & Speech* 43, 377–401.

Matras, Y., & Schiff, L. (2005). Spoken Israeli Hebrew revisited: Structures and variation. *Studia Semitica: Journal of Semitic Studies Supplement*, 16, 145-191.

Neuhauser, S., and A. P. Simpson. (2007). “Imitated or Authentic? Listeners’ Judgements of Foreign Accents.” In *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS)*, 1805–1808.

Nolan, F., & Asu, E. L. (2009). The pairwise variability index and coexisting rhythms in language. *Phonetica*, 66(1-2), 64-77.

Podesva, R. J. (2011). Salience and the social meaning of declarative contours: Three case studies of gay professionals. *Journal of English Linguistics*, 39(3), 233–264.

Pratt, T. and A. D’Onofrio. (2017). Jaw setting and the California vowel shift in parodic performances. *Language in Society* 46: 283-312.

Preston, D. R. (1996). Whaddayaknow?: The modes of folk linguistic awareness. *Language*

*awareness*, 5(1), 40-74.

Preston, D. R. (2013). Language with an attitude. *The handbook of language variation and change*, (pp. 157–182).

Purnell, T., Idsardi, W., & Baugh, J. (1999). Perceptual and phonetic experiments on American English dialect identification. *Journal of language and social psychology*, 18(1), 10–30.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>

Ramus, F., & Mehler, J. (1999). Language identification with suprasegmental cues: A study based on speech resynthesis. *The Journal of the Acoustical Society of America*, 105(1), 512-521.

Ramus, F., Nespors, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 73(3), 265–292.

Sasson-Levy, O. (2013). A different kind of whiteness: Marking and unmarking of social boundaries in the construction of hegemonic ethnicity. In *Sociological Forum*, vol. 28, (pp. 27–50). Wiley Online Library.

Schilling-Estes, N. (1998). Investigating ‘self-conscious’ speech: The performance register in Ocracoke English. *Language in Society* 27: 53-83.

Sclafani, J. (2009). Martha Stewart behaving badly: Parody and the symbolic meaning of style. *Journal of Sociolinguistics* 13: 613–33.

Silber-Varod, V., Cohen, E., Cohen, E., Strull, I., & Aharonson, V. (2022). Towards a public segment catalogue of spoken Hebrew. URL:

[https://www.researchgate.net/publication/360463146\\_Towards\\_a\\_Public\\_Segment\\_Catalogue\\_of\\_Spoken\\_Hebrew](https://www.researchgate.net/publication/360463146_Towards_a_Public_Segment_Catalogue_of_Spoken_Hebrew)

Silverstein, M. (2003). Indexical order and the dialectics of sociolinguistic life. *Language & communication*, 23(3-4), 193–229.

Smith, B. L., Wasowicz, J., & Preston, J. (1987). Temporal characteristics of the speech of normal elderly adults. *Journal of Speech, Language, and Hearing Research*, 30(4), 522-529.

- Swirski, S. (1989). *Israel: The oriental majority*. Zed books.
- Torgersen, E. N., & Szakay, A. (2012). An investigation of speech rhythm in London English. *Lingua*, 122(7), 822-840.
- Tzfadia, E., & Yiftachel, O. (2008). Sfariferia: the development towns and the mizrahi space [Hebrew: Sfariferia: Arej ha-pituax ve-hamakom ha-mizrahi]. *Block: Architecture/ Media/ Theory/ City Journal* (6), 22–30.
- Vaughn C., Kendall T. (2019). Stylistically coherent variants: Cognitive representation of social meaning/Variantes estilisticamente coerentes: Representação cognitiva de significados sociais. *Revista de estudos da linguagem*, 27(4), 1787–1830.
- Vicenik, C., & Sundara, M. (2013). The role of intonation in language and dialect discrimination by adults. *Journal of Phonetics*, 41(5), 297-306.
- Yaeger-Dror, M. (1988). The influence of changing group vitality on convergence toward a dominant linguistic norm: An Israeli example. *Language & Communication*, 8(3-4), 285–305.
- Young, N. (2021). The sociolectal and stylistic variability of Rhythm in Stockholm. *Language and Speech* 65(4), 1034-1070.
- Zlotnick, E., Dzikiewicz, A. J., & Bar-Anan, Y. (2015). Minno.js (version 0.3). [Computer software].