



**THÈSE DE DOCTORAT**  
**DE L'UNIVERSITÉ PSL**

Préparée à l'École Normale Supérieure

**Hearing ethnicity:  
Classification, stereotypization and processing of  
socially marked phonetic features in Modern Hebrew**

Soutenue par

**Si BERREBI**

Le 16 Décembre 2021

Ecole doctorale n° 540

**Lettres, Arts, Sciences  
humaines et sociales**

Spécialité

**Sciences cognitives**



Composition du jury :

Erez LEVON  
Professeur, University of Bern

*Président  
Rapporteur*

Shiri LEV-ARI  
Maîtresse de conférence,  
Royal Holloway University of  
London

*Rapportrice*

Noam FAUST  
Maître de conférence,  
Université Paris 8

*Examineur*

Roey GAFTER  
Maître de conférence,  
Ben Gurion University of the  
Negev

*Examineur*

Outi BAT-EL  
Professeure, Tel-Aviv University

*Co-directrice de thèse*

Sharon PEPERKAMP  
DR, École normale supérieure

*Co-directrice de thèse*

## RÉSUMÉ

---

Cette thèse explore la relation entre la perception sociale et la perception phonologique. Le contexte social est la dichotomie fondée sur l'ethnie entre les juifs israéliens “Mizrahi” (d'origine moyen-orientale ou d'Afrique du nord) et “Ashkenazi” (d'origine européenne).

La partie I étudie les effets de la variation phonologique et du discours qui l'entoure, sur la perception sociale : c'est-à-dire les effets de ses productions sur la perception de l'identité du locuteur par les auditeurs. Les principaux résultats empiriques de cette étude sont les suivants : (i) L'opinion selon laquelle la variation parmi les jeunes locuteurs natifs n'est pas conditionnée par l'ethnie est affectée par un biais idéologique qui considère l'une des identités comme étant “générale”. (ii) Les auditeurs ont perçu de jeunes locuteurs natifs ayant une phonologie similaire comme étant significativement différents les uns des autres en termes d'ethnicité, sur la base d'un enregistrement de ~5 secondes (Expérience 1); (iii) Le principal corrélat acoustique des jugements des auditeurs était le rythme de la parole (un %V plus faible était corrélé avec une “européanité” perçue).

La partie II étudie les façons dont l'expérience de la variation phonologique affecte le traitement phonologique en ligne. Le chapitre 3 révèle que les locuteurs qui fusionnent [h] et [χ] en [χ] en production peuvent encore avoir un encodage lexical séparé de ces phonèmes dans le lexique, à condition qu'ils aient été exposés à la parole non-fusionnée depuis l'enfance. En particulier, dans une tâche de décision lexicale (expérience 2), les auditeurs qui fusionnent mais qui ont été exposés à la parole non-fusionnée depuis l'enfance ont rejeté les mots à catégorie commutée [χ]→[h] à un taux significativement plus élevé, indiquant que les locuteurs de la même variété (c'est-à-dire les locuteurs qui fusionnent) peuvent encore différer significativement dans leurs représentations lexicales. L'expérience 3 montre que les auditeurs qui fusionnent réussissent bien à distinguer entre [h] et [χ], ce qui indique que le modèle divergent de l'expérience 2 est le résultat d'une différence dans l'encodage lexical, et non dans la catégorisation phonétique. Il a été démontré que le personnage d'un locuteur enregistré jouait également un rôle dans le traitement : les réponses aux éléments fusionnés [h]→[χ] étaient médiées par le personnage du locuteur, de sorte que les items fusionnés produits par un locuteur typique de la variante avec fusion étaient acceptés plus souvent et faisaient l'objet d'une réponse plus rapide par rapport aux items fusionnés produits par un locuteur typique de la variante sans fusion.

Les expériences 4-5 (chapitre 4) se concentrent plus spécifiquement sur le mécanisme de traitement orienté vers le locuteur, en utilisant des variantes marquées qui sont partagées par au moins deux groupes sociaux. Les résultats ont révélé un effet du traitement orienté vers le locuteur : lorsque la différence acoustique entre les catégories était subtile, l'exposition au personnage connu pour produire la catégorie la moins connue a renforcé la catégorisation phonétique, comme le montrent les taux de précision et les temps de réaction aux mêmes stimuli enregistrés.

Pris ensemble, les résultats des expériences 2-5 fournissent une double contribution à notre compréhension du traitement de la parole dans la voie indicielle : (i) Une connaissance profonde d'une langue vernaculaire peut induire des prédictions subtiles concernant la production ; même

lorsque la forme produite n'est pas incorrecte au niveau global, mais seulement incompatible avec l'identité du locuteur, elle était rejetée plus souvent ou traitée plus lentement par des auditeurs très expérimentés. (ii) Certaines variantes linguistiques n'ont pas une connotation aussi forte que d'autres ; présenter un personnage qui utilise de telles variantes peut focaliser l'attention des auditeurs sur la dimension acoustique pertinente.

## MOTS CLÉS

---

Perception phonologique, Sociophonétique, Sociolinguistique, Reconnaissance auditive des mots

## ABSTRACT

---

This thesis explores the relationship between social and phonological perception. The social setting is the ethnically-based dichotomy between "Mizrahi" (Middle Eastern background) and "Ashkenazi" (European background) Jewish-Israelis.

Part I studies the implications of phonological variation and the discourse around it on social perception; i.e., the implications of one's productions on listeners' perception of their identity. The main empirical findings of this investigation are: (i) The view that variation among young native speakers is not ethnically-conditioned is affected by an ideological bias that views one identity as 'general'. (ii) Listeners perceived young native speakers with similar phonology to sound significantly different from each other in terms of ethnicity, based on a ~5 seconds long recording (Experiment 1). (iii) The main acoustic correlate of listeners' judgments was the rhythm of speech (lower %V was correlated with perceived Europeanness).

Part II investigates the ways in which experience with phonological variation affects online phonological processing. Chapter 3 reveals that speakers who merge [ħ] and [χ] into [χ] ("Merged speakers") in production may still have separate lexical encoding of these categories in the lexicon, provided they have been exposed to unmerged speech since childhood. Particularly, in a lexical decision task (Experiment 3), merged speakers exposed to unmerged speech since childhood rejected switched-category [χ]→[ħ] words at a significantly higher rate, indicating that speakers of the same variety (i.e. Merged speakers) may still significantly differ in their lexical representations. Experiment 2 shows that Merged listeners of all linguistic backgrounds perform well in distinguishing between [ħ] and [χ], indicating that the divergent pattern in Experiment 2 is the result of a difference in lexical encoding, not in phonetic categorization. The persona of the recorded speaker was shown to play a role in processing as well: responses to merged [ħ]→[χ] items were mediated by the persona of the speaker, such that merged items performed by the 'Merged' speaker were accepted more often and responded to faster compared with merged items performed by the 'Unmerged' speaker.

Experiments 4-5 (Chapter 4) focus more specifically on the mechanism of speaker-oriented processing, utilizing marked variants that are shared between at least two social groups. The results revealed another effect of speaker-oriented processing: when the acoustic difference between categories was subtle, exposure to the persona that is known to produce the less-known category bootstrapped phonetic categorization, as reflected in accuracy rates and reaction times to the same recorded stimuli.

Taken together, the results of Experiments 2-5 contribute to our understanding of speech processing in the indexical (social) route the following insights: (i) Deep acquaintanceship with a vernacular can induce subtle predictions regarding production; even when the produced form is not illicit at the global (lexical) level, but only inconsistent with the identity of the speaker, it was rejected or responded to more slowly by highly experienced listeners (Experiment 3). (ii) Some linguistic variants are not as salient as others; presenting a persona who uses such variants can focus listeners' attention on the relevant acoustic dimension (Experiment 5).

## KEYWORDS

---

Phonological perception, sociophonetics, sociolinguistics, auditory word recognition

## Table of Contents

<b>Acknowledgments</b>	<b>v</b>
<b>List of Figures</b>	<b>1</b>
<b>List of Tables</b>	<b>2</b>
<b>I Social meanings of phonological variation in Modern Hebrew</b>	<b>4</b>
<b>1: Approaches to linguistic variation in the history of research on Modern Hebrew</b>	<b>5</b>
1.1 Modern Hebrew as a tool of national demarcation . . . . .	6
1.2 First impressions of variation . . . . .	9
1.2.1 “Language planning” . . . . .	9
1.2.2 New ethnic categories, new linguistic markers . . . . .	12
1.3 Variation among native speakers of Modern Hebrew . . . . .	15
1.3.1 The alleged convergence towards a “Hebrew Koiné” . . . . .	15
1.3.2 Why do Mizrahis converge to the Ashkenazi spoken variety? . . . . .	18
1.3.3 Individual-level “convergence” . . . . .	19
1.4 Variation and “the periphery” . . . . .	20
1.5 Variation and socio-economic status . . . . .	22
1.6 Third-wave sociolinguistics: identity construction in a segregated space . . . . .	25
1.7 Conclusions and overview of the thesis . . . . .	27
<b>2: Perception of Ethno-linguistic phonological variation in Modern Hebrew</b>	<b>32</b>
2.1 Previous studies of ethno-linguistic phonological variation in MH . . . . .	33
2.1.1 Quantitative production studies . . . . .	33

2.1.2	Qualitative studies and impressionistic observations . . . . .	36
2.1.3	Perception studies . . . . .	40
2.2	Experiment 1: Ethnic classification in ~5 seconds . . . . .	44
2.2.1	Methods . . . . .	44
2.2.2	Results and discussion . . . . .	50
2.2.2.1	Inter-listener agreement . . . . .	50
2.2.2.2	Acoustic correlates of the ratings . . . . .	53
2.2.2.3	The influence of listeners' linguistic background and attitudes . . . . .	54
2.3	Conclusions . . . . .	55

## **II Exploring phonological representations of marked variants in Modern Hebrew** **58**

<b>3:</b>	<b>Pharyngeal minds: different lexical representations for speakers of a single dialect</b>	<b>59</b>
3.1	Introduction . . . . .	59
3.1.1	Sociolinguistic background . . . . .	60
3.1.2	Perception of merged categories . . . . .	62
3.2	Experiment 2: Discrimination between [h] and [χ] in speakers with merged productions . . . . .	64
3.2.1	Methods . . . . .	64
3.2.1.1	Participants . . . . .	64
3.2.1.2	Stimuli . . . . .	64
3.2.1.3	Data analysis . . . . .	65
3.2.2	Results . . . . .	65
3.2.3	Discussion . . . . .	65
3.3	Experiment 3: A switched-category lexical decision task . . . . .	66
3.3.1	Methods . . . . .	67
3.3.1.1	Participants . . . . .	67
3.3.1.2	Stimuli . . . . .	68
3.3.1.3	Procedure . . . . .	69

3.3.1.4	Data analysis . . . . .	69
3.3.2	Results . . . . .	70
3.3.2.1	Acceptance rates . . . . .	70
3.3.2.2	Reaction times . . . . .	71
3.3.3	Discussion . . . . .	72
3.4	General discussion and conclusions . . . . .	74
<b>4:</b>	<b>Exploring phonological variation with the IAT: Advantages and limitations</b>	<b>77</b>
4.1	Background: speaker-oriented predictions . . . . .	78
4.2	The Implicit Association Task in context . . . . .	79
4.3	Experiment 4: [h]~[χ] and [e]~[ej] . . . . .	83
4.3.1	Experiment 4a: “Broadcaster” ([h]+[ej]) . . . . .	84
4.3.2	Experiment 4b: “Yemenite” ([h]+[e]) . . . . .	86
4.3.3	Success rates in categorization of target sounds in Experiments 4a+b . . .	87
4.3.4	Discussion . . . . .	89
4.4	Experiment 5: [h]~[x] and [r] . . . . .	91
4.4.1	Experiment 5a: “Palestinian” . . . . .	94
4.4.2	Experiment 5b: “Argentinian” . . . . .	95
4.4.3	Success rates in categorization of target sounds in Experiments 5a+b . . .	95
4.4.4	Discussion . . . . .	101
4.5	General discussion . . . . .	102
4.6	Conclusions . . . . .	105
<b>5:</b>	<b>Conclusions</b>	<b>106</b>
5.1	Implications for variation studies on Modern Hebrew . . . . .	106
5.2	Implications for phonological theory . . . . .	107
5.3	Implications for online speech processing in the indexical route . . . . .	109
	<b>References</b>	<b>112</b>
<b>Appendix A:</b>	<b>Screenshots from an online debate on “Mizrahi speech”</b>	<b>126</b>

<b>Appendix B:</b>	<b>Full models of Experiments 2-3</b>	<b>132</b>
<b>Appendix C:</b>	<b>Full models of Experiments 4-5</b>	<b>135</b>



## Acknowledgments

I am primarily indebted to my advisors, Outi Bat-El and Sharon Peperkamp. The guidance and inspiration Outi provided me from *Beginners Phonology* to this submission has shaped my understanding of language and linguistics and provided me with a broad perspective for which I am ever thankful. Her ideas are engraved in my way of thinking about phonology.

The extreme privilege of working with Sharon at the LSCP was likewise formative. I thank her for giving me a second home in Paris, for the endless consultations and comprehensive support which was at all times both professional and personal to the highest possible standards.

To my reviewers Erez Levon and Shiri Lev-Ari, I am grateful for insightful feedback on every aspect of this work – from the methodological, through the statistical, to various theoretical aspects. I truly enjoyed our discussion at the defense, which got me excited about possible follow ups. I also thank Heather Burnett and Alejandrina Cristia for providing helpful feedback on earlier stages of the project.

To Almog Simchon, my brother of choice and expert of psychology and statistics in need, I am grateful for hours of careful thought and close assistance during one of the most complicated periods of time. This work could not have been carried out without your support.

Chapter 1 had three major inspirations: Yaakov Ben Tolila's 1984 dissertation analysis of phonological variation in 'Romema', which provided the main insight on the dynamics of 'convergence' in Modern Hebrew; and works by Shlomo Morag and Yehudit Henshke, who carefully documented ethnically conditioned variation in the traditions of Yemen and the Maghreb, respectively. I also thank Zvi Ben-Dor Benite for helpful feedback and an illuminating discussion of the analysis, and Yaara Benger Alaluf for many conversations on early 20th century Palestine, as well as referring me to relevant works beyond the immediate fields of linguistics and sociology.

Chapter 2 has benefited from consultations about phonetic variation in Hebrew with Michael Becker, Roy Becker-Kristal and Pavel Ozerov. I am grateful to Yoav Bar-Anan and Elad Zlotnik, the developers of Minno, for their generous support, going as far as developing new tools needed for the implementation of the experimental designs presented in Chapters 2 and 4. Finally, I thank the many participants from friends circles, university students, and particularly the members of the Facebook group 'Jadaijot – medabrot Feminism Mizrahi', for their thoughts and feedback.

I am grateful to my collaborators in Chapter 3, Noa Bassel and Roey Gafter, for pushing my ideas forward and assisting me in many crucial steps along the way. This project began as a term paper in Roey's seminar on language and ethnic identity in Israel, and he has contributed his thoughts and feedback on other chapters as well, at all stages of the work. Noa has been my partner on this from the start, and we'll hopefully continue to bounce ideas off each other for as long as

we live. Every insight in this project is at least partly due to her. I also thank Guy Shoshani and Alon Fishman, who recorded the stimuli for this chapter; my students at the BGU and the Hebrew University, as well as friends and family members, for participating; and audiences at the Atelier de Phonologie in University Paris-Xiii, the Hebrew University linguistics departmental seminar, the Tel Aviv University Cognition and Language Learning lab (CaLL), NWAV 2021, and the Bar Ilan University linguistics departmental seminar.

I thank Zizo Abul Hawa, Daniel Hadad, Noam Faust and Yaakov Tzuberi for recording stimuli for Chapter 4. In addition to this technical help, Noam has also been a supportive friend I could always turn to, for helpful arguments about phonology and for good times at Buttes Chaumont.

To my high-school music teachers Igor Verhovsky and Alexander Vardi, I thank for showing me that a good teacher provides at the same time a shortcut and inspiration to keep pursuing the road. I thank my teachers at Tel Aviv University and the Hebrew University, Elitzur Bar-Asher Siegal, Edit Doron, Naama Friedmann, Aya Meltzer-Asscher, Malka Rapaport-Hovav and Tal Siloni, for plenty of shortcuts and inspiration. I am particularly thankful to Roni Katzir for his perspective on what's interesting about language, for the questions he helped me consolidate and for helping me pursue them through encouragement and feedback.

To my friends and colleagues in Paris, Tel Aviv and Jerusalem: Inbal Abarbanel, Odelia Ahdout, Aviad Albert, Amir Anvari, Bar Avineri, Monica Barbir, Adi Behar Medrano, Melissa Berthet, Noa Brandel, Megan Dailey, Hila Davidovitch, Milica Denic, Ittamar Erb, Ava Guez, Naomi Havron, Yuval Katz, Maayan Keshev, Nikki Koesterich, Georgia Loukatou, Mora Maldonado, Eyal Marco, Shirly Orr, Laurel Perkins, Andrés Soria Ruiz, Hualin Xiao, Suma Samara, Camila Scaff, Yechezkel Shabanov, Aviv schoenfeld, Hadas Yeverechyahu – thank you for the pleasure of learning with and from you. I am especially thankful to Hadas Zaidenberg, who taught me how to teach and love it; Daniel Asherov for phono-marathons in great company; Ezer Rasin for comfort, support, and challenges; and Tal Ness and Doron Veltzer for crucial technical help, always patiently and sometimes on a very short notice.

To the research students forum of Academia for Equality, Smadar Ben Natan, Mor Cohen, Efrat Ben Shoshan Gazit, Shai Gortler, Nitzan Tal, Revital Madar, Maayan Ravid – thanks for being a place for sharing and diffusing academic and bureaucratic anxiety; and to Radhia Achheb, Shira Batash-Galili, Dafna Lebowitz, Merav Saroosi-Peretz, Eric Sinaman and Ruti Zussman, thanks for helping me through this bureaucracy with patience. To Halah Abdelhade and Dana Molcho, thanks for providing support from the distance.

I thank Miriam Bassel, Yael Hadad and the entire Bassel family for having made the most meaningful contribution to chapter 3, and for many happy moments and unconditional support. Finally, I thank my family, and especially my sister Shir, and my parents Miri and Eli, for providing a home with a profound sense of grounding and love that I carry with me everywhere.

## List of Figures

2.1	An example of the *HLH% contour . . . . .	39
2.2	An example of the *LH% rising contour . . . . .	39
2.3	%V by intonation phrase (left to right: I-III), ethnic background and age. . . . .	47
2.4	Hierarchical clustering of perceived ethnicity . . . . .	51
3.1	Results of Experiment 2. . . . .	65
3.2	A summary of the logistic regression model of acceptance rates in Experiment 3 .	71
3.3	A summary of the linear regression model for RTs in Experiment 3 . . . . .	73
4.1	The pictures used in the presentation of the speakers . . . . .	85
4.2	Distributions of D-scores as a function of speakers' guise in Experiment 4. . . . .	90
4.3	Distributions of D-scores as a function of speakers' guise . . . . .	96
4.4	A visual description of the error rates model, Experiment 5 . . . . .	98
4.5	A visual description of the RTs model, Experiment 5 . . . . .	100
A.1	Post in Politically Corret, a feminist facebook page . . . . .	127
A.2	Post in Ha-'arsim ba'im, a Mizrahi activists Facebook page . . . . .	128
A.3	Continuation of the post in Ha-'arsim ba'im from A.2 . . . . .	129
A.4	"There is no accent" . . . . .	129
A.5	"No body shaming and no accent" . . . . .	130
A.6	"Using an accent is not necessarily racist" . . . . .	130
A.7	"This bothered me as well, the sketches made by these cartoonists are always racist" 130	
A.8	"Thank you for bringing this up, the racism and body shaming bothered me as well" 131	

## List of Tables

2.1	Markers of Mizrahi speech and their corresponding letters . . . . .	35
2.2	Results of Devens 1981 . . . . .	42
2.3	Summary of proposed socially meaningful phonetic variables in Modern Hebrew . . . . .	43
2.4	Self-reported demographic data of participants in Experiment 1 . . . . .	45
2.5	Personal background of recorded speakers by ethnic background and age . . . . .	46
2.6	Acoustic features of the recordings used in Experiment 1 . . . . .	46
2.7	Mean rating and SD by cluster, age and gender . . . . .	52
2.8	Model output predicting ethnicity ratings for older speakers in Experiment 1 . . . . .	54
2.9	Correlation between the Spread of Ratings and views regarding variation. . . . .	55
3.1	The three target groups of listeners in the study. . . . .	60
3.2	Merged and non-merged realizations. . . . .	62
3.3	Prediction for pattern of acceptance of switched-category stimuli . . . . .	67
3.4	Sample stimuli in the Merged and Pharyngealizing voices . . . . .	69
4.1	The distribution of marked variants in guises in Experiment 4 . . . . .	77
4.2	An example of the IAT. . . . .	80
4.3	Examples of stimuli in Experiment 4 . . . . .	84
4.4	Block order and examples of stimuli in Experiments 4a-b. . . . .	86
4.5	Results by Stimulus Type and Guise in the practice blocks . . . . .	89
4.6	Results by guise, stimulus and association in target blocks, Experiment 4 . . . . .	89
4.7	The alignment of features according to guise in Experiment 5. . . . .	92
4.8	The procedure and sample stimuli of the SC-IAT in Experiment 5. . . . .	94
4.9	Results of Experiment 5 . . . . .	97
4.10	Overt judgments of participants in Experiment 5 . . . . .	102
B.1	Accuracy model of Experiment 2 . . . . .	132

B.2	Reaction times model of Experiment 2 . . . . .	133
B.3	Acceptance rates model of Experiment 3 . . . . .	133
B.4	Reaction times model of Experiment 3 . . . . .	134
C.1	Reaction times model of the practice blocks of Experiment 4 . . . . .	135
C.2	Error rate model of the practice blocks of Experiment 4 . . . . .	136
C.3	Summary statistics of the generalized linear mixed model of Experiment 5 . . . . .	137
C.4	Full Reaction times model of Experiment 5 . . . . .	138

**Part I**

**Social meanings of phonological  
variation  
in Modern Hebrew**

# I

## Approaches to linguistic variation in the history of research on Modern Hebrew

Israel has a population of about 9.2 million, out of which 73.9% are Jewish and 21.1% Palestinian Arab (Muslims, Christians and Druze).<sup>1</sup> The primary languages spoken in the country are Modern Hebrew (approximately 5 million speakers), Arabic (1.4 million) and Russian (1 million).

Modern Hebrew has emerged as a spoken variant of the Hebrew language at the end of the 19th century, following waves of immigration to Israel, then Ottoman, followed by British-occupied (Mandatory) Palestine. Immigration persisted throughout the 20th century, rendering an array of languages from the origin countries as contact-languages of Modern Hebrew, including various dialects of Arabic, Ladino, Yiddish, Aramaic and Russian. Consequently, linguistic variation among speakers of Modern Hebrew is in many cases closely related with their ethnic background. The connection between ethnic identity and linguistic variation is generally not reflected in much of the linguistic literature on MH, where a narrative of unidirectional standardization is dominant. In supporting this narrative, linguistic research on Modern Hebrew follows a longstanding tradition in the political discourse surrounding the language, which, to the extent that it recognizes variation, presents it as a fault to overcome.

The current chapter surveys approaches to inter-speaker variation in the linguistics literature on Modern Hebrew, and shows that in most cases (though not all), it points to a general reduction in ethnically-conditioned variation. I highlight data that suggest otherwise, at every historical stage.

I argue that current attitudes towards linguistic variation in Modern Hebrew in the literature as well as the general public, and in particular a popular attitude of *denial* of the existence of ethnically-conditioned variation outside of three markers that are extensively discussed in the literature, is affected by earlier ideological structures. The first significant ideology is that in order to turn Hebrew into the native language of Jews in Ottoman and then Mandatory Palestine, the space should be clear of other languages (Section 1.1). Then, when it comes to variation within Modern Hebrew, two particularly strong ideologies are that of language planning, which was supposed to ensure that the most “authentic” and “aesthetic” language would be disseminated to speakers

---

<sup>1</sup>This excludes the Palestinian population of the West Bank and Gaza, refugees and immigrants (about 200,000), but includes the settlements in the West Bank and the Golan heights. Data are drawn from the Israeli Central Bureau of Statistics, 2020 report, available on their website [www.cbs.gov.il](http://www.cbs.gov.il).

(see 1.2); and the “Melting Pot” ideology-turned-policy, which strove to a complete levelling of variation within the new native community, reflecting the unity of the nation (Section 1.3). These ideologies have obscured the state of affairs of variation as well as the processes of the reduction in variation. I argue that the levelling of variation is not a general phenomenon that spread evenly throughout the native Hebrew speakers community. Instead, based on historical documentation of the spread of speakers with different linguistic backgrounds within the state of Israel and their degree of segregation, I argue that variation was retained and developed in small secluded communities relying on linguistic features from the first generation’s L1 Arabic substrata. At the same time, convergence occurred many times, at many stages, at the level of the individual speaker who wishes to avoid stereotyped variants. In other words, the process of levelling is ongoing - children who live in segregated communities typically face pressure to abandon stereotyped sociolinguistic markers when leaving their small community. This individual-level convergence hypothesis also fares better in explaining the current situation of variation in MH: if levelling began in the 1950s, why isn’t the process completed by now?

Section 1.4 discusses the correlation between variation and place of residence. Section 1.5 reviews variation studies conducted from an educational and developmental perspective, in which consistent divergences from a prescriptive norm were attributed to individual deficiencies. In Section 1.6 I review an ethnography in which variation in MH is taken to be the result of the speaker’s personal choice for identity construction, and show in which cases this perspective can contribute to our understanding and in which cases it might obscure it.

I claim and demonstrate that the variety that arises from variation research, the “Mizrahi/ Peripheral” dialect of Hebrew (labels that will be unpacked through this chapter), requires more documentation than it has received so far, and special care should be given to the source of variation, as well as to the source of reduction in variation.

### **1.1 Modern Hebrew as a tool of national demarcation**

The emerging variants of Modern Hebrew were idolized by most Zionist leaders in the beginning of the 20th century as a linguistic reflection of the Jewish national identity. In the following decades, this ideological link between language and national identity reshaped social divisions within the multicultural society of Palestine.<sup>2</sup>

Historical sources describe a surge of development and urbanization in major cities at the end of the 19th century, accompanied by diverse immigration from the Ottoman empire, including Arabic-speaking Jewish communities that were immersed in the local Muslim-majority society (Gross 1976; Kark 1984a). In this context, the new waves of Jewish immigration from Europe stood out in maintaining the use of European languages, mostly German, Yiddish and Russian. This is the context in which the Zionist leadership made its first efforts to establish Modern Hebrew

---

<sup>2</sup>The discussion in this section owes much to the analysis presented in Bengier Alaluf and Bassel (2021).



as the only language of the Jewish community. The focus on Hebrew was also meant to establish new social borders, along national rather than geographical, social and cultural axes.

As Irvine and Gal (2000) demonstrate, language differentiation is an important tool in creating social boundaries. Irvine and Gal (2000:38) detect three semiotic processes used for language differentiation: **Iconization** is the projection of semiotic features into personal features; **Erasure** – the elimination of data that challenge the ideology; and **Fractal recursivity** – the projection of an opposition at some level of representation onto inter-group relations. Similar processes have also shaped the linguistic discourse of the Jewish Zionist community. I argue that in Zionist thinking, Hebrew has undergone a process of iconization, making it a symbol of national sovereignty.

The role of language in establishing social boundaries is emphasized in the motivations underlying the foundation of Tel Aviv next to Jaffa. At the beginning of the 20th century, the Jaffa area was a multi-cultural and multi-linguistic society, with immigrants from Egypt, Syria, Lebanon, Afghanistan and north Africa flowing to the city, including Jewish immigrants. The Jewish Ottoman community of this period integrated in the city life (Kark 1984b), and was later joined by waves of European-Jewish immigration. Between 1880 and 1909, twenty new neighborhoods had been established, eleven of which were Jewish (Rotbard and Gat 2015). At this point, Jaffa was seen as the center of culture and commerce by all of its residents, who maintained inter-group ties.

The foundation of Tel Aviv marked a change of approach in some of the Jewish residence, as the following brochure for investors (1906) demonstrates:

96% of the Jewish inhabitants of Jaffa live in Arab houses, and the rent they pay every year is no less than 40,000 francs. This money ends up in the pockets of strangers [*zarim*, lit. ‘stranger’ or ‘foreigner’], fortifying them and destroying us at the same time. If we can get this sum in our hands every year, it will take little time before we acquire huge sums. That is why we need to buy a large swath of land as soon as possible, on which we will build our houses. It should be positioned next to Jaffa, and it will become the first Hebrew city, where 100% Ivrim [lit. ‘Hebrews’] will reside, *where Hebrew will be spoken*, where cleanliness and purity will be kept, and we shall not go in the ways of the gentiles (cited in Yekutieli-Cohen (2010):6, my translation and emphasis)

This quote illustrates that the crucial motivation in the founding of Tel Aviv was to define a social boundary, based among other things on the use of Hebrew. Having a “hygienic” city, with Hebrew as the *only* language on the street, is a goal in and of itself. Importantly, during this period it was not the case that all Jewish immigrants spoke Hebrew among themselves. In 1904, only 6 out of 29 Jewish schools in the country taught in Hebrew; the first Jewish academic institution in the country, the Technion, planned in 1913 to open its gates and teach in German. The Technion’s decision to be a German-speaking institution had started a heated debate, in which students from the first Hebrew-speaking school, Gimnasja Herzlia (‘Herzl Gymnasium’, founded in Jaffa in 1905 and moved to Tel Aviv in 1909), protested with their teachers against the decision.

The protest, which resulted in diminished funding for the Technion, eventually worked, and the institution decided to teach in Hebrew. Graduates of the Gimasja later founded the Battalion for the Defence of the Language, a group against the use of languages other than Hebrew by Jews, that operated primarily in Tel Aviv. They spread pamphlets that demonstrate “common mistakes” which originate in Yiddish, and promoted the presence of Hebrew in public spaces, by among other things tearing down shop signs that included languages other than Hebrew and demanding the municipality to proofread signs in Hebrew that contained errors (Zuckermann 2020). In a book titled “fun Nyu-York bis Rehovot un tsurik” *‘from New York to Rehovot and back’*, the Yiddish poet Solomon Blumgarten (pen name: Yehoash, a Biblical Hebrew name) wrote about Tel Aviv in 1914:

“Yiddish here is more defiled than a pig. Speaking Yiddish in the street takes a great deal of courage.”

The role of language in the planning of social boundaries was further reflected in debates of the Tel Aviv municipality welfare department during the 1930s-40s, regarding social relations between Jewish women and Arab men. Razi (2011) finds that the protocols of the department raise concerns for Jewish girls and women who lived in mixed neighborhoods and were integrated in the local society. The command of Arabic by Jewish youth is specifically mentioned as a risk factor, that enables Jewish youths to establish undesired contact with Arab Jaffa residents. The Arabic speaking Jewish women in particular were viewed as crossing national and ethnic boundaries, and measures have been taken to “rehabilitate” them, including cooperation of the secular social department with the Rabbinic institution of Tel Aviv, which established the “committee for guarding the honor of the daughters of Israel” in 1942. These debates highlight the role assigned to language by Zionist authorities, in breaking existing ties between Jews and Arabs and re-positioning residents of the area on the national-religious axis.

The segregation created by the enforcement of geographic (Tel Aviv) and national (Jewish) linguistic boundaries was complemented by acts of force, including reporting on Arab presence in the street of Tel Aviv as a hazard, incidents of detentions of Arab workers in Tel Aviv facilities, and finally in the occupation of the Jaffa metro-area in a series of military operations. During 1948, most of the Palestinian residents of Jaffa were expelled and from a community of over 85,000 only 3,647 residents remained under martial law in the neighborhood of Ajami (Monterescu 2019 and references there). Following these actions, the goals of Tel Aviv’s founders in turning the multi-cultural metropolitan of Jaffa into a largely monolingual, uni-national space were achieved.

Under this mindset, variation among Hebrew speakers, which was inevitable from day one due to the diverse backgrounds and contact languages in the Hebrew speaking community, was perceived as challenging to the newly formed social division. Variation along the cultural/linguistic axis seemed to contain a potential to preserve or encourage unwanted divisions, e.g. Arabs and Jews from Middle Eastern backgrounds as opposed to Jews immigrating from Europe. In addition, language was perceived as an emblem of the resurrection of the nation – a flag of sorts – and

variation was viewed as a disturbance, an error that does not reflect the nation properly. The following section shows how these attitudes are reflected in early debates on variation in Modern Hebrew.

## **1.2 First impressions of variation**

### *1.2.1 “Language planning”*

Among the linguistic resources available to speakers of Modern Hebrew were diverse literary and liturgical traditions, which include prose and poetry, theological and philosophical debates and the extensive use of the language for communication as a lingua-franca of the Jewish world. This state of affairs naturally invites variation across periods and geographical areas, as argued in Morag (1990). The emergence of the new spoken variant is commonly referred to as ‘The Hebrew Revival’ (e.g. Haršav 1993, among many others), a term which in itself erases any process of change prior to the end of the 19th century. Thus, Morag offers the alternative term ‘full return to Hebrew’ (1990:74).

The spoken varieties of Hebrew were affected by the linguistic traditions of the different ethnic groups within the community. The liturgical reading traditions of Hebrew can be roughly divided into Ashkenazi (of European descent), Sepharadi (of Middle Eastern descent) and Yemenite, which have distinct reading traditions that have spread through the different Jewish communities. In addition to variation arising from these traditions, speakers of Hebrew were also native speakers of other languages, that also had various contact languages.

During this period, variation seems to have been referred to in public debates mainly as an obstacle, with reference to strategies which may be used to reduce it. This approach to the Modern Hebrew language, as something that needs to be shaped and then disseminated to the people, has been rather stable in the public discourse from the full return to Hebrew: native speakers of the language often testify that they “don’t speak Hebrew well”, a linguistic insecurity that is complemented by institutions that publish, once in a while, “instructions” as to how the language should be spoken.

The authors I follow in the current section are Zionist thinkers who acquired Hebrew as a second language. Their varied approaches to the way Hebrew should be used, and more specifically *pronounced*, reflect two main language ideologies, related to the same process of iconization discussed in Section 1.1, that Hebrew is a living symbol of national sovereignty. One is an ambition that Hebrew should be authentic (i.e. respecting its biblical past) and aesthetic. This ideology usually surfaces as a concern that the language would become inauthentic or unaesthetic because of its careless speakers. In prescriptive discussions of variation, whether centered on aesthetics or authenticity, the need for a structured method is emphasized. A seemingly contradictory ideal, that emerged simultaneously, is the superiority of the *native* speaker as the hero of the Hebrew ‘revival’. The native speaker was idolized by Eliezer Ben Yehuda (who is often credited for turning Hebrew into a spoken language) since, as he acknowledges, despite his efforts to think only in

Hebrew (and of course, speak and write only in the language), when his thoughts wander, they are often organized in Yiddish, with borrowed words from Russian or French (Haršav 1993).<sup>3</sup>

Thus, two agents of linguistic prestige arose during this period: the language itself, due to its role as a cultural practice and national symbol that needs protection; and at the same time, the native speaker - the “Tsabar”, lit. ‘cactus’ - a sign of renewed authenticity and nativity (Mor and Sichel 2015).

In a booklet entitled “the Hebrew Accent” (1930), Ze’ev Jabotinsky, a leading Zionist writer, exemplifies the first trend as he outlines the features of the desirable Hebrew accent:

“The language is the core and basis of the national tune. As the violinist or pianist works on a sonata that he will tomorrow play in public, so should everyone work on improving their accent [...] Without shame, I confess that the ‘*taste*’ [ta’am; also “stress”, S.B] at the basis of the scheme proposed in this booklet is a European taste, not an ‘eastern’ one. The reader will find a clear trend in my proposals to break free of those speech sounds that have no fulcrum in the phonetics of Western languages - a clear tendency to bring our accent as closely as possible to the received musical ideal of Europe; the same ideal, the same musical perspective by which, for instance, the Italian language is considered beautiful and the Chinese language not beautiful. This is the perspective I chose, first of all since we are Europeans and our musical taste is European, the taste of Rubinstein and Mendelssohn and Bizet. But also from the objective side of the problem I am certain, based on the arguments I put forth above, that the accent I propose is indeed closer to the “correct” accent, to the ancient tune of our language in the mouths of our ancestors, than the Arabic pharyngeal imitation; and certainly more than the offhanded, lawless, tasteless accent with which we vernacularized our speech and humiliated our language, one of the greatest and most noble languages in the world, to a level of toneless noise without character” (Jabotinsky 1954(1930):115-116; my translation)

Jabotinsky emphasized the need to pick the most “aesthetic” variants of pronunciation, and he suggests that every speaker should be active in their attempt to make Hebrew as “beautiful” (in his view, central/western European) as possible. It is worth mentioning that he dislikes not only the sounds associated with Arabic and with the liturgical traditions of Sephardi and Yemenite Jews. He also scolds Ashkenazi speakers for their tendency to delete schwa in word initial unstressed positions (e.g. [kneset], cf. [kəneset], ‘parliament’). His booklet often gives “notes” to native speakers of a particular European language - Russian and German (whose speakers should beware of devoicing a word final obstruent, a process that shouldn’t occur in Hebrew in his view), English (whose speakers should note that there are no lax vowels in Hebrew). Still, his persistence on the

---

<sup>3</sup>Recent analyses of the full return to Hebrew acknowledge that the actual role of single agents cannot be as vast as previously described, and the role of educators in rural areas should receive more attention (e.g., Haršav 1993; Nahir 1998; Matras and Schiff 2005). Still, Ben Yehuda is widely praised in Israel in schools, books, a popular song and even a musical called (*Bustan Spharadi* lit. ‘a Sephardi orchard’), as “The reviver of Hebrew”.

need to eradicate characteristics of speech that emphasize the Semitic origins of Hebrew exceeds these additional concerns.

Jabotinsky's proposals, e.g. to eliminate the pharyngeals, were far from accepted. Most scholars who were invested in the idea of making Hebrew the native language of the Jewish people believed that these sounds were part of the received pronunciation of ancient Hebrew, and that maintaining the Semitic vibe of the language should be a priority. The positive attitude towards the pharyngeals and [r], the apical trill produces in the Sephardi and Yemenite traditions, was made public in the documents and decisions of the Hebrew Academy and by its predecessor, *Va'ad ha-Lashon ha-'Ivrit* 'the Hebrew language committee'.<sup>4</sup> One of its founding members, David Yellin, who was also the founder of the first Hebrew College for Teachers, personally responded to Jabotinsky's critique of Hebrew pronunciation. The conjecture that Hebrew is closer to Western languages than to Arabic was, to Yellin, ridiculous:

The author has an issue with life, with Jewish people of different origin countries who speak different languages, that affect their pronunciation when they speak Hebrew. He notes many wrongs [*kilkulim*, lit. 'wrong pl.' but also 'spoiled', 'broken', S.B] in the accent, and advises the reader as to how to avoid them. The advice comes as orders and "notices". A central issue seriously devalues it [Jabotinsky's review, S.B]; it is the author's opinion that the Hebrew language is not primarily oriental, and one should not mention its affinity with the Arabic language and its pronunciation. [Allegedly, S.B] The land of Israel has been the home of many peoples, among whom descendants of Western countries, and the Hebrew language has been influenced by their speech more than it did by the east. There are many paradoxes in the world, let this be another one. This conjecture is the author's personal, subjective opinion; his aesthetic taste leans towards European-ness, and all that is accepted among Europeans is the absolute good that we should all hold. Of course, in matters of taste there is no point arguing, but we cannot subject the science and history to our own taste." (Yellin 1930:149-150; my translation)

Although Yellin presents a much more sympathetic attitude towards the pharyngeals, the framework remains the same: speakers are expected to adopt linguistic behaviours dictated by experts and taught in school. He even agrees with Jabotinsky on some pronunciation points:

His comments on the pronunciation of the strong emphasis [a geminate, S.B], that we should revive it in order to give color and power to our speech, are indeed true, and so are his comments about the Schwa mobile.

---

<sup>4</sup>The Hebrew language committee later became the Hebrew Academy, an institute that to this day regulates the usage of Hebrew in public, including in the education system, in governmental documents and most importantly in the Israeli broadcasting authority, that operates radio stations and a TV channel. Its authority is regulated by the 1953 law of "the Supreme Institution for the Study of the Hebrew Language".

The debate on whether Hebrew should sound more Sepharadi or more Ashkenazi is already very much alive at this point. The more interesting part is that unlike later discussion, the pharyngeals and [r] are not the only features on the table. Instead, both Jabotinsky and Yellin discuss a variety of ethnically-marked linguistic features: geminates and the emphatic consonants (that were produced by some native Arabic speakers), stress (that was mostly penultimate in Ashkenazi varieties and final in the other traditions), schwa deletion (mostly in Ashkenazi varieties), the production of the historically long vowel [e:] (produced as [ej] by some Ashkenazi speakers).

Prescriptive discussions of this rich array of linguistic features had quickly died out, although it is not clear at which point variation was levelled; indeed, in the case of [e]~[ej], variation still exists (see Chapter 2). In public discourse, variation on various dimensions other than the pharyngeals and [r] are still mentioned well into the 1970s. For example, in an ethnicity classification perceptual experiment (Devens 1981), listeners insisted that they can distinguish between L1 Arabic Hebrew speakers of different dialectal backgrounds, e.g. listeners described Moroccans as having “harsh sibilants”, Iraqis as most likely to produce the emphatic [q], and Yemenis as having a particular ‘color’ to the pharyngeal fricative.<sup>5</sup>

These claims have not been investigated qualitatively. The distinction between pharyngealizers and non-pharyngealizers was deemed socially important, while within-group distinctions were not. These sounds were common to Yemenite and Sepharadi speakers of various origin languages (including various dialects of Arabic, but also Aramaic and Farsi), who were now grouped under the umbrella term “Mizrahi” (lit. ‘oriental’).

### 1.2.2 *New ethnic categories, new linguistic markers*

The crystallizing ethnic boundaries between Jewish Israelis shaped the discourse on variation in Modern Hebrew. In some cases, it also shaped variation itself. Below, I describe two cases in which variation was shaped by these newly-formed boundaries: erasure of variation within the diverse “Mizrahi” category, and the conscious choice of the Sepharadi stress system by Ashkenazi speakers, that led to reduction of variation within the group of native speakers. Before introducing these cases, a note is required about the ethnic categories themselves.

Mizrahi and Ashkenazi are cover terms for various ethnic identities (Khazoom 1999, Shohat 2001, Gafter 2016a). The etymology of the terms *Mizrahi* and *Ashkenazi* is not symmetric. ‘Ashkenaz’ is the historical name of an area of Germany, originally used to refer to a dominant Jewish community in west Germany that has its own liturgical tradition and religious practices. The comparable term to Ashkenazi Jews was *Sepharadi* ‘Spanish’ Jews, which referred to the Jewish community of Iberia that developed separated religious practices and traditions. Both communities spread to other areas, either physically (by migration) or culturally (by religious and philosophical influences): Ashkenazis to eastern and western Europe, and Sepharadis to North Africa and the Middle East.

---

<sup>5</sup> There is no evidence that any of these impressions contributed to actual ethnic identification. More details on this study are found in Chapter 2.

The “Ashkenazi” and “Sephardi” labels are still used to denote religious practices, alongside other traditions such as the Yemeni and Ethiopian ones, but Ashkenazi remained the ethnic definition of European Jews, while Sephardi was substituted with *Mizrahi*, literally ‘oriental’, as the ethnic category of Jewish communities in Muslim countries (i.e. this is a broader definition, that covers Yemenis as well, although they have a separate liturgical tradition). The geographical dimension is also irrelevant for this category boundary - many Mizrahis came from countries to the west of Israel, like Morocco, Algeria, Tunisia and Libya, while some Ashkenazis lived to the east of Israel (e.g. Russia). The origin of this usage instead stems from a history of orientalizing: a process in which one group uses the dichotomy of East/West to present another group as degraded (Goffman 1963; Said 1978; Khazoom 1999). In fact, a similar usage of the idea of the orient (which was coined in German – “Ostjuden”), preceded the massive immigration waves from North Africa; it was originally used by German Jews to denote their eastern European counterparts from Poland, Russia, the northern parts of Hungary and Romania. Rather than depicting geographic relations, ‘eastern-ness’ in both of these incarnations alludes to the establishment of this category as an opposition to “Western Culture” (Khazoom 1999). This is an example of fractal recursivity (Irvine and Gal 2000): the “eastern” category, associated with primitiveness and poverty, applies within the society time and time again to mark marginalized identities. There are also processes of erasure here: “Mizrahi” speakers have come from diverse backgrounds, and more importantly, from different linguistic backgrounds. Their native language (different varieties of spoken Arabic, but also Farsi, for example) and liturgical traditions (Yemeni and Sephardi) must have influenced their speech in different manners, just as Jabotinsky describes for speakers of different European languages. However, most of this variation was not mentioned by researchers, and all of these speakers were classified as belonging to the same ethnic group in most variation studies. The erasure of “within-group” variation further consolidated the categories of “Koine/General” and “Oriental” Hebrew, as they were later referred to in the literature reviewed in Section 1.3.

Modern Hebrew ended up using the Sephardi vowels and final stress pattern (Morag 1990; Zuckermann 2005). The fact that the Sephardi stress pattern prevailed is somewhat unexpected. According to Even-Zohar (1979), the insistence on the Sephardi stress pattern was a form of rejection of the Ashkenazi accent, as a marked foreign variety that belonged to the Old World and the Nomad Jew, who was depicted as weak, wandering the world without a connection to the land, not sovereign. It might seem mysterious that the stress pattern, and not, for example, segmental features, were chosen to mark the social boundary between these two fictitious personae; after all, in previous discussions, segmental features were discussed at least as much. Phonological theory sheds light on this point: Yiddish is a stress language, in which stress is manifested in pitch accent and syllable duration, just as in Hebrew (Lötzsch 1990, cited in Svetozarova et al. 1999). Also, stress in Yiddish is affected by morphological structure and there are exceptions to the generally penultimate pattern, such that it must be memorized in certain cases (Jacobs 2005:135). Thus, Yiddish speakers were aware of stress and able to manipulate it, making it a minimal yet salient feature that can be used to distinguish the veteran immigrants from the new stream of immigrants and refugees from Europe. On the other hand, the segmental features of the Sephardi system

were not available to Yiddish speakers, who traditionally merged [ħ] and [χ] and omitted [ʕ] (not to mention the emphatic consonants and geminates). As I show in Chapter 3 for speakers of a current merged variety of Hebrew, while these speakers can distinguish between [ħ] and [χ] phonetically, they were unable to represent them as separate categories in their lexicon. I assume that the same was true for the Yiddish speakers from whom they inherited this merger.

Within a semiotic analysis of the linguistic situation of Modern Hebrew, the adoption of the Sephardi stress pattern is a case of iconization: the penultimate stress pattern was chosen to represent the Yiddish-speaking Old Jew, thus allowing speakers who wish to set themselves apart from this persona to borrow the Sephardi (largely ultimate) stress pattern.

The most salient linguistic dimension of ethnic variation, both in public discourse and in the literature, was the use of the pharyngeal consonants [ħ] and [ʕ], as well as the coronal tap/trill [r]. These markers had an ambiguous social meaning: on the one hand, the prescriptive norm was to keep these sounds, as a conscious attempt to maintain an authentic Semitic vibe to the language, and suit its orthographic system that represents each of these categories as a letter. To Sephardi and Yemenite speakers these sounds were natural to produce because of their presence in Arabic and in their liturgical Hebrew practices. On the other hand, European speakers – who formed most elite groups within Israeli society, as described below – did not pronounce the pharyngeals. Their phonological strategy was to collapse these sounds with their uvular counterpart [χ]; or total elimination in the case of [ʔ] and [ʕ]. Despite the prescriptive norm, then, speakers were motivated not to produce the pharyngeal sounds, that were quickly associated by the elite with the Orient and every stereotype Europeans pinned on it during the colonialist period (Davis 1984; Shohat 2001). Following Irvine and Gal (2000), I argue that the pharyngeal sounds were involved in the following semiotic processes:

1. Iconization: The pharyngeal sounds were first perceived as markers of “the Orient” and at the same time as representatives of the “Semitic vibe” that the language should have. Based on the social status of their speakers, described below in sections 1.3 and 1.5, they evolved to be perceived as markers of poor education and resistance to modernity.
2. Erasure: The pharyngeal sounds were stationed as the ultimate divide between Jews from Muslim countries and their European counterparts. The label *Mizrahi* (lit. ‘eastern’) was adopted as a cover term for the former. Any other source for diversity across different communities was overlooked.
3. Fractal recursivity: The Zionist administration attempted to unite the various Jewish communities under one national and linguistic identity, while at the same time preserving European race ideologies and reconstructing them within the Jewish population, from a dichotomy between Jewish and Arab-Muslim, to European-Jewish vs. Arab-Jewish. The pharyngeal sounds were used as gatekeepers to this ends (e.g. Khazoom (1999); Shohat (2001)).



The following section shows that much of the linguistic research on Modern Hebrew took an active role in consolidating these processes.

### **1.3 Variation among native speakers of Modern Hebrew**

The efforts to shape citizens' cultural and linguistic preferences that is evident in the writings of early Zionists had turned, with the founding of the state of Israel, part of a comprehensive policy set to ensure the integration of Jewish immigrants, which was termed by Israel's first prime minister, David Ben Gurion, 'the Melting Pot'.<sup>6</sup> The Zionist administration strove to high involvement in citizens' lives, including forcing immigrants to take up a Hebrew name (Azriyahu 2000), the founding of a centralized education system that resisted parental involvement in education (Zameret 2012), enforcing mandatory youth programs with the national army Israel Defence Force (IDF), and mandatory military service at the age of 18 for both men and women.

The declared goal of the melting pot doctrine was to eliminate cultural diversity among the Jewish community, including linguistic variation. "Hebrew culture", "Hebrew music", "Hebrew youth", were all coinages used to emphasize the new desirable identity (Shavit 1984; Sheleg 2010; Peleg 2018) Accordingly, early linguistic research on Modern Hebrew generally focused on processes of standardization and levelling in spoken Hebrew. In this section, I review these studies and propose an alternative to the received view, according to which a Hebrew has undergone a fast and steady process of standardization.

#### *1.3.1 The alleged convergence towards a "Hebrew Koiné"*

Blanc (1968) describes a "koinization process" towards a national standard, within which he distinguishes four styles or registers on an axis of formality. When doing so, he describes four generations of native Hebrew speakers and stylistic variation between them, which is expressed in sentences constructed by the author. In other words, this fairly canonical study of Modern Hebrew is not based on linguistic or sociological data. Blanc describes the first speakers of Modern Hebrew as "overwhelmingly raised on East European (Ashkenazi) variety of traditional pronunciation" (p.243), ignoring the influences of Yemeni and Sepharadi liturgical traditions on speakers who were taught Hebrew from scripture from a young age (e.g. Morag 1963, 1990).

Regarding the phonological inventory, Blanc describes two variants of Hebrew: 'Oriental Hebrew' and 'General Hebrew'. The main distinctive feature between them is the usage of pharyngeal sounds. According to his report, all Mizrahi Jews over 40 speak 'Oriental', that is, with pharyngeals, while the 'general' system is spreading at younger ages. On the other hand, Blanc states that General Hebrew might include pharyngeal sounds as a *stylistic* feature. That is, the pharyngeal

---

<sup>6</sup>The same term has a separate yet related history in the United States, but in the case of Israel it had become an official governmental doctrine.

sounds characterize Oriental Hebrew, unless used by an Ashkenazi speaker – in which case they are ‘stylish’.

No data is provided as evidence of the existence of a general dialect of spoken Hebrew, nor to the extent to which it is widespread. General Israeli is described as “the common core to which new speakers (children, immigrants) tend to be approximate”. It is further claimed that General Israeli is spreading among speakers of Mizrahi heritage, again without data. Instead, the author provides his personal impression that Mizrahi speakers over the age of 40 invariably speak “oriental”, while younger speakers converge towards General Israeli. This discussion erases any variation between Mizrahi speakers of different communities, as well as between various Ashkenazi communities, and claims for a national standard which is not well-defined or studied.

Another theme in this paper is that of Europeization, which is a term given by the author to the influences of European contact languages. He argues that such influences are widespread in General Israeli and that they are not perceived as markers of ethnicity, but rather as markers of Modernity and Western values. Other contact languages, e.g. Arabic, are not mentioned.

Evidence that this analysis is far from capturing the full picture of variation in Modern Hebrew of the time are found in a governmental meeting protocol from roughly the same years. In 1971, a protest group named the Israeli Black Panthers was responsible for a wave of demonstrations against systemic discrimination towards Mizrahis in the state. The protocol of their only meeting with then prime minister Golda Meir reveals the extent to which language variation was dominant in the Hebrew speaking society. The meeting begins with an apology:

PM Meir: Are you all from Jerusalem?

Reuven Abergel: Yes. We beg your pardon for our simple language, we cannot express ourselves in literary Hebrew, but only as we were taught. (in Lev (2008), p. 198, my translation)

PM Meir continues to interrogate Abergel regarding his family and himself:

PM Meir: How many children are there in your family?

Reuven Abergel: We were fourteen, and ten of us are still alive.

Meir: Are you the eldest?

Abergel: No, I’m in the middle. My older brothers are married.

Meir: Do they work?

Abergel: They too work in despicable occupations. First they were in criminal institutions. A big part of us has become degenerate and has no urge to work. No one trained us for hard work. *My brother has been living in this country for 25 years, and he does not know how to speak Hebrew. I speak better, because I’m in the streets more often.* He’s illiterate, like the rest of my brothers. All that - since we have been getting terrible education [...]. (in Lev (2008):199, my translation and emphasis)

The description indicates the 20-year-old speaker's strong awareness to the contrast between his Hebrew and the language spoken in the halls of the government, as well as his internalization of the perception of his speech as a marker of cultural subordination.

In the two citations above, Abergel argues that his incompetence in Hebrew is due to educational gaps. He stresses that even being in the streets, he has become more proficient in Hebrew than he and his brother had become in school. That is, segregation was at its peak during the years in which Blanc conducted his study, and given the great turnouts in Panthers' demonstrations, it is likely that Abergel's description of the variety his family members speak represents other young people living in segregated neighborhoods. Next, Sa'adia Marciano, another panthers' leader, protests Meir's line of questioning that hints that the Panthers and their family members are not doing enough to get themselves out of poverty:

Marciano: You keep asking us why we don't work. Will they let us study? If they will, our youth will go and work of their own free will. This is one of the things I think are necessary: I am as old as this state and I was born here, *and still I have a Moroccan accent*. I think this teaches us several things.

PM Meir: I've been in this country for 50 years and people tell me that I have an American accent to this day. This doesn't teach us anything.

Marciano: Did you acquire education in the United States?

PM Meir: Of course. (in Lev 2008:206, my translation and emphasis)

While Marciano takes accent to be a proxy for being educated within the mainstream of the native environment, Meir prefers to remain at the surface level, and point out that she too has an accent, i.e. that accents are not a real problem for integration. Marciano attempts to clear his point again - she was brought up in the US and received her education there, therefore it makes sense that she should have an American accent - but the PM seems not to understand, or to purposely not engage with his message. In any event, the patronizing attitude of the PM continues throughout the meeting, where she advises Marciano "as I would a son or a grandson" in her words, to start acquiring an occupation, work and study, so that he could perform as a role model for his brothers and future children.

This conversation indicates that the announced "melting pot" policy has not been as over-throwing as presented. The quoted speakers have immigrated to Israel at the ages of four and one year old, yet they report speaking in an ethnically marked variety and experiencing an educational neglect that is unexpected if the melting pot agenda was to be applied broadly within the Jewish population. To the contrary, the most powerful tools in standardization of language, which is equality in education and desegregation in residency, are the opposite of the highly-segregated Zionist education system, which contributed to variation rather than reducing it (e.g., Dahan and Yona 2013).

The conception in Blanc's study, that "most speakers" are aligned to an emerging standard leaves behind the greater part of the population, as argued below. This is particularly important since Blanc's account was highly influential to any future studies on variation in MH. In fact, I could not find a paper on the topic that does not cite him, mostly building on his claim that variation is in steady decline.

### 1.3.2 *Why do Mizrahis converge to the Ashkenazi spoken variety?*

The literature continued to report on an ongoing process of standardization and levelling among Hebrew speakers. Yaeger-Dror (1988, 1993) lays out a similar picture to Blanc (though with far more data), using the parallel terms 'Mizrahi Hebrew' and 'Modern Koiné' for 'Oriental' and 'General' respectively. Yaeger-Dror argues that Mizrahi Israelis lose the pharyngeal segments due to processes of integration and their growing "group vitality". She relates the decrease in usage of pharyngeal sounds with an upgrade in the social status of Mizrahis. Instead of arguing whether a causal link can be drawn between these two processes, I pose the question: has the socioeconomic status of Mizrahi Jews in Israel improved in the decades parallel to their loss of pharyngeals?

This question obviously exceeds the scope of the current work, but consider Cohen (1998), that found, in a survey of statistical data between 1975 and 1995, that the gaps in education level between Mizrahis and Ashkenazis in the relevant period have decreased by about 2%, while the gap in income rate between the groups has expanded by at least 20%.

It is additionally perplexing that a *rise* in "group vitality" is predicted by Yaeger-Dror to induce *fewer* productions that are characteristic of that social group. It seems to contradict the way social prestige is translated into linguistic traits in many of the cases described in the literature: an increase in prestige of a social group goes hand in hand with an increase in the rate of production of the dialect associated with prestige (Labov 1972; Trudgill 1972; Eckert 2000, 2008b; Lev-Ari and Peperkamp 2014, among many others).

Yaeger-Dror further suggests that anti-Arab sentiments that characterize Mizrahi speakers contribute to the loss of linguistic features that are shared with Arabic, a claim that is echoed in many other works in the social sciences (e.g., Mizrahi and Herzog 2012). This claim is somewhat contradicted by the retention of other cultural customs associated with an Arab identity: while the production of pharyngeal sounds was considerably reduced, Henna ceremonies, loan words from Arabic, Arabic style in fashion and Middle Eastern cuisines are still very widespread among the younger generations of Mizrahi communities. Under these circumstances, it is somewhat unpredictable that the pharyngeal sounds decreased in usage so fast while being used by first generation speakers and having the support of the writing system.

A different view is found in Shlomo Morag's works (1963; 1990, a.o.) on Modern Hebrew in general, and the Hebrew of Yemenite Jews in particular. Morag describes the Sephardi Hebrew and the Yemenite Hebrew as a cultural legacy rather than an inferiority marker (1990:89). He argues that these communities have demonstrated advanced education systems, and had less

linguistic immigration struggles due to their developed Hebrew skills. The fact that their dialect was in retreat, Morag attributes to the effects of social stratification in the Jewish society, under which the political, cultural and social elite members were of an East European descent, and the remainder split between Central Europeans, Western Europeans, and Palestine-born Jews (among whom some were of Middle Eastern descent, but their immigration took place around eight generations before). Below, I outline a more elaborate proposal as to how such processes of retreat from the native variety occurred in practice.

### 1.3.3 *Individual-level “convergence”*

The conflict between the prescriptive prestige of Mizrahi sounds and their low prestige when actually used by Mizrahis has led to unusual sociolinguistic phenomena. A notable example is the linguistic variant used by the media for a number of decades. In governmental and military radio stations and television - the only legal stations and also the most popular until commercial television started in the 1990s - broadcasters were instructed to use the three marked phonemes of the Mizrahi inventory, which was prescriptively “correct”, but broadcasters were also expected to maintain certain Ashkenazi features of speech, such as the production of the diphthong [ej], thus creating a new, highly recognizable style (Ben David 1974).

The ambiguous status of the pharyngeals is also demonstrated in public views of variation. Davis (1984) cites a survey conducted in the 1980's, according to which 70% of the participating students from Haifa University responded that either Yemenis or Iraqis, ethnic groups included within the social construct “Mizrahi” and associated with preserving the marked Mizrahi inventory, speak “the most correct Hebrew”.

As Ben Tolila (2002) notes, “in the Israeli socio-linguistic market, the value of Mizrahi Hebrew [here, meaning the use of the pharyngeals and [r]; S.B] is either too low or (in rare cases) too high. In either case it is perceived as less desirable” (my translation, p.133). He bases this claim on data collected for a quantitative and qualitative variation study during the 1970s, on a small farmers community who have all immigrated to Israel from the same village in Morocco. The recordings included 46 speakers in four speaking styles (free conversation, interview, paragraph reading, word list) based on the method introduced in Labov (1972), that has previously showed that variants might be used differently depending on the level of formality.

This corpus is particularly valuable, first due to its size – it is the largest tagged corpus of Mizrahi speakers in the periphery from that period of time; and second, since Ben Tolila's analysis takes into account social factors that have not been considered before. Most notable is the attention given to power relations between large families, that have existed for years in Morocco before the immigration to Israel but were reshaped upon arrival by contact with new communities, mainly the majority-Ashkenazi administrative authorities. This recognition of local patterns of variation and social differentiation are lacking from the majority of variation studies on Hebrew; Blanc and Yaeger-Dror, for instance, discuss global effects of “prestige”, that is attributed to the

“General” variety of Modern Hebrew. Fleshing out the effects of social organization on individual productions acknowledges the histories of immigrants and their social organization. Another often-neglected social factor that Ben Tolila (2002) considers is the political agenda of speakers, that ranges between a desire to assimilate in an Ashkenazi-majority environment, introduced usually at the regional high-school (the community was otherwise ethnically homogeneous, including the local elementary school); and a desire to be accepted as a local leader, which may have more complex relations with the chosen productions, since being relatable is crucial for a social organizer.

Ben Tolila (2002) shows interactions of these factors with the chosen production of the pharyngeals: at kindergarten and elementary school, most of the speakers in his sample use pharyngeals; in high school, most produce the Ashkenazi variant instead; and among university graduates, all produce the Ashkenazi variant. It should be noted that the sample is very small in terms of speakers – only two university graduates are recorded – and yet, the trend in the data makes it seem likely that the older speakers in the sample have changed their production at some point, after having more intense relations with Ashkenazi speakers, starting from the regional high-school.

I hypothesize that the “reduction of variation” that Blanc and Yaeger-Dror describe as a general, consistent process is composed of various case studies like that of the small town described in Ben Tolila (2002): variation depended on gender (in his sample, Israel-born women above elementary school all produced the Ashkenazi variant; a similar relationship between gender and “global” prestige was found in many other cases, beginning with Trudgill 1972); on the amount of daily contact with Ashkenazis (the only Israel-born men who used the Ashkenazi variant in an open conversation worked in the municipality); and on family membership (members of one of the two dominant families in the village were more likely to use the Ashkenazi variant). Variation was in fact not overall reduced, but began to accumulate new meanings and pattern differently based on the speaker’s social background and their interlocutors.

According to my hypothesis, the ethnic homogeneity of residential areas, as well as schools, has had a crucial role in the retention of variation. Such “pockets”, in which the influence of an Arabic substratum is stronger, still exist today. On the other hand, the stereotypization of features associated with Mizrahi speech has contributed to the loss of the variants most associated with the variety: the pharyngeals and [r]. In the next section, I support the claim of ethnic homogeneity during the formative years of the state based on sociological and geographic studies.

#### **1.4 Variation and “the periphery”**

As described above, features from the Arabic substratum tended to be reduced in Ashkenazi-majority areas and retained in Mizrahi-majority areas. That is why Henshke (2013a; 2013b; 2013c) refers to the Hebrew vernacular influenced by Arabic as “peripheral Hebrew”, which ties Mizrahi speech to a (not well-defined) geographical area. The usage of the label “periphery” in Israel is different from its original geographic meaning and requires some elaboration.

Originally, the borrowing of the word “periphery” into Hebrew denoted geographical distance from the centre of the country, indicating areas with reduced accessibility in terms of infrastructure and transportation, which leads to reduced accessibility to job markets, education and services - resulting in a predictable correlation with low socio-economic status (hereafter SES). However, when it turned out that geographic parameters are inferior to ethnic parameters as predictors of socio-economic status, the term “socio-economic periphery” was put forth to indicate low income areas regardless of location. Thus, the “periphery” was broadened to include low income residences in the center of Israel, while excluding high income residences in the geographic periphery.

In a technical list, the Israeli government defines “the social periphery” such that it covers most small towns and villages in the south of Israel as well as parts of the north, not covering Kibbutzim (lit. “gathering, clustering”, collective communities founded by socialist Ashkenazis from the beginning of the 20th century) and high income villages in the same areas. Cities in the center are generally not included, but every city has neighborhoods on this map, specified to the level of streets. All of these locations are residences of non-Ashkenazi majority communities. According to this list, the Israeli periphery amounts to 31% of the country, including neighborhoods at the very center of the Tel Aviv-Jaffa metropolitan area.<sup>7</sup> In order to understand how this complex picture of geographic, social and economic stratification came about, one has to recall the first architectural plan for Israel, written between 1949-1952 (Sharon 1951).

The idea behind the plan was to repopulate the villages from which Palestinians were expelled, in order to prevent their return and fill the space with new Jewish immigrants (Yiftachel 2021:96-102 and references there). Yiftachel (2021) identifies four key components of the plan, that shape the spacial and economic development of Israel to date. The two features most relevant to our linguistic interests are that the plan relied on establishing many new settlements, mostly on or near the ruins of emptied Palestinian villages, in order not to leave lands abandoned for the refugees; and on separation between Jews and Palestinians and Jewish ethno-classes (roughly during those years, Mizrahi and Ashkenazi). In the rural areas, many Kibbutzim (populated by an overwhelming majority of Ashkenazis) were already established. In order to be accepted, newcomers have to pass an interview with a committee of Kibbutz members. Development towns, on the other hand, were populated almost exclusively by Mizrahis (Tzfadia and Yiftachel 2008). They were later joined by immigrants from the USSR and Ethiopia.

Ethnic homogeneity was not limited to rural areas. Hasson and Gonen (1974) studied the patterns of residential mobility in immigrant housing estates at the outskirts of two medium-sized towns at the center of the state, Kfar Sabba and Rishon lezion. The cities were established prior to the state of Israel, and their center was populated by veteran Ashkenazi communities. In the outskirts of the two cities, as in many other towns during the 1950s, immigrant absorption camps (“ma’abarot”) were set in order to receive the newcomers. In Kfar Sabba and Rishon lezion, most of the immigrants were from the Middle East, but there was also a minority from Europe. The

---

<sup>7</sup>The full list of residences included in the governmental definition of “social periphery” is available at: <https://negev-galil.gov.il/media/2253/list2021.pdf> [Hebrew].

researchers found that in the years following the process of turning the camps into neighborhoods, most of the Europeans had left the outskirts and moved into the Ashkenazi-majority center of the city, or into newly built neighborhoods, making the area even more ethnically homogeneous.<sup>8</sup>

Given these facts, it was noted by sociologists, activists and comedians that the term “periphery” is a euphemism for communities of Palestinian, Mizrahi, Ethiopian and ex-USSR majority; as put by sociologist Nissim Mizrahi: “the periphery in Israel is a matter of demography, not geography”.<sup>9</sup> The “peripheral” dialect can be viewed as a mix of the two: an initial population of (mostly) Arabic speakers learning the same target language in a relatively segregated area, joined by subsequent waves of immigration with additional substrata, that learn the target language along with the Arabic substratum.

### **1.5 Variation and socio-economic status**

As the protocol of the meeting between PM Meir and representatives of the Israeli Black Panthers demonstrated (Section 1.3.1), gaps in formal education between Mizrahis and Ashkenzis were robust from the beginning: despite the declared aspiration on forming a united Jewish identity, the Israeli government took on a highly segregated approach in all walks of life. Mizrahi Jews were populated in different towns and geographical areas than Ashkenazi Jews, were educated in a separate school system run by the ministry of labor, sent to different tracks in their (mandatory) military service etc. This policy was also expressed in an unequal allocation of land (Yiftachel 2021), which has led to a high correlation between ethnicity and socioeconomic status, that persists to date (Cohen 1998; Roby 2015). This state of affairs is not unusual in situations of language variation, and requires taking means to control the interaction of the two factors. In the Israeli case, the general approach in most studies was to completely reduce ethnic variation to socioeconomic gaps.

The educational literature on linguistic variation is particularly persistent in collapsing the ethnic and socioeconomic dimensions. For example, in a large-scale quantitative analysis of phonological and morphological variation in various ages of Modern Hebrew speakers, Schwarzwald (1981) defines the social categories of speakers as “Middle class” (high SES) and “culturally disadvantaged” (in Hebrew “te’unej tipu’ax”, low SES), without specifying the speakers’ ethnic backgrounds. The second label requires some elaboration: according to the works of Karl Frankenstein (1905-1990), a leading Zionist educator whose writings are still taught in teachers’ seminars in Israel, the educational system must bridge over deficiencies and gaps that a child developed during their first years of life. The term “culturally disadvantaged”, borrowed from

---

<sup>8</sup> The areas of Rishon Letsion described in Hasson and Gonen (1974) as Mizrahi-majority include the neighborhood Ramat Eliyahu, to date the poorest neighborhood in the city, also specified as “periphery” in the technical list of the government. When the study came out, the neighborhood was populated by a majority of Mizrahi and Romanian immigrants, who were joined (some replaced) in the 1990s by immigrants from the former USSR and Ethiopia.

<sup>9</sup> <https://www.calcalist.co.il/local/articles/0,7340,L-3726214,00.html> [Hebrew]



studies in English pedagogy, has been used by the ministry of education to mark mainly poor children from Mizrahi families. Their alleged incompetence had been used as a justification to put them in special classes and even special schools for culturally disadvantaged children, causing de facto segregation within those cities that included Ashkenazi as well as Mizrahi residents (Swirski 1991).<sup>10</sup> The “care” scale (Hebrew “madad ha-tipu’ax”) of the ministry of education is still used to describe the average socio-economic level of children in a school, and different tracks are still mostly populated along the ethnic axis: vocational tracks for a majority of non-Ashkenazi children, and academic tracks for a majority of Ashkenazi children. Descriptions of the cognitive and linguistic deficiencies of “disadvantaged” children can still be found in contemporary literature on education, despite the understanding that the category is clearly ethnically marked. In a lecture to students from 2011, a professor at the department for Hebrew studies at the University of Haifa said: “In the linguistic aspect, many fundamental deficiencies were found in this group: inferiority in abstraction and generalization, difficulty in categorization. Studies found that the language of a culturally disadvantaged child is deficient in its quality and quantity. It is not merely the result of mental backwardness, but also of social backwardness” (Fisherman 2011; my translation).<sup>11</sup>

When Schwarzwald (1981) uses the term “culturally disadvantaged”, she also conceals the ethnic identities of speakers in each group. Still, there are several indices that allow the reader to deduce who in particular she means. In the chapter that describes the variables and the groups who participated in the experiments, she acknowledges that language varieties may diverge from each other in a non-hierarchical manner, citing Labov (p. 55). However, she takes a completely different approach in the analysis: instead of investigating the language from *within*, contrasting structures and productions as speakers of a given vernacular perceive them and relying on systematicity as the main characteristic of a vernacular, she compares both groups to a single prescriptive norm – the imagined variant described and disseminated by the Hebrew Academy. Yet, even the prescriptive standard does not hold in all cases. When it comes to the pharyngeals, Schwarzwald (1981) decides not to consider underlying pharyngeals that are produced in the merged Ashkenazi variety as “mistakes”:

Performance of [ʕ] as [Ø] and the merged variety of [ħ] as [χ], as well as failure to produce geminates, were not considered mistakes. (Schwarzwald (1981), p.56, my translation)

That is, variation is highlighted when the prescriptive norms are congruent with the productions of the strong social group, and erased from the analysis when prescriptive norms side with the weak social group. Unfortunately, there are no other large-scale analyses of this sort (320 participants) that can tell us about the state of variation in the production of the pharyngeals in different age

---

<sup>10</sup> Later, immigrants from the USSR received a similar treatment.

<sup>11</sup> The situation of Arab-Palestinian schools within the Israeli educational system was far more difficult. At the first years of the state, Arab citizens lived under martial law. Arab children were completely segregated and their schools received a dramatically lower budget. Ironically, they had to fight to get the label of “culturally disadvantaged”, which gave (meager) financial benefits. They also fought to get the vocational schools to open in their villages as well. Only in the 2000s Arab students started to be divided into tracks (Swirski 1991, 2018).

groups during the 1970s. Relying on Ben Tolila's (1984) study from roughly the same years, much depends on the degree of segregation from Ashkenazi communities.

In 1995, Ravid presented another large-scale variation study with an educational and psycholinguistic perspective, using a method of experimental elicitation with 188 speakers of several age groups. This study also reduces the social categories into non-ethnic ones, using the similar socioeconomic categories of "disadvantaged"/ "uneducated"/ "low SES" and "literate"/ "high SES". Still, Ravid does not ignore ethnicity altogether. She clearly states that the low SES group is Mizrahi, and the high SES group is Ashkenazi (p. 30), but maintains that the main difference between the groups is educational, citing the well-known and steady gap in education between Mizrahis and Ashkenazis. Since most of the study is dedicated to opaque phonological and morphological alternations, it makes sense to attribute the amount of linguistic input, usually correlated with income (Hart and Risley 2003), to be of primary importance. Still, in order to exclude the possible influence of the speaker's linguistic background - particularly, the Mizrahi sociolect - it requires a serious control group of adults. If adults overwhelmingly choose the "erroneous" form, it isn't erroneous, just different (as Ravid acknowledges in the discussion, but not in the conclusions). The study includes more than twice the number of high SES Ashkenazi than low SES Mizrahi participants (58 vs. 130), only 17 of them are adults, some of whom not native speakers of Hebrew. Indeed, these speakers showed more "errors" than the other groups, indicating that younger Mizrahis might not just have difficulty in conjugating words, but rather rely on forms that they have previously heard in their close environment.

Despite these data, Ravid (1995) assumes that there are no local dialects of Modern Hebrew, because "Israel is too small and Modern Hebrew too young" to develop local dialects (1995:8). Both parts of this argument are weak: first, stating that Modern Hebrew is too young ignores the development of Hebrew separately prior to its full return, in liturgical and literary standards that evolved in different communities, under the influence of various contact languages. Many speakers arrived with partial or even full command of the language as readers and writers, which is likely to have influenced their speech (Ravid is aware of all that, as evident from the introduction section about Hebrew). While acknowledging that under the standard for "native language" in cognitive science, Hebrew is rather young, the influences of knowledge from literary standards and from L2 is known to induce variation (e.g., Khamis-Dakwar et al. 2012). Second, claiming that Israel is too small ignores the fact that segregation and social differentiation, and not only geographic distance, are the necessary conditions for the development of variation. Israel is indeed small, but as mentioned above, it is highly segregated. It is also worth mentioning that at the same geographic area, Palestinian Arabic includes several distinct dialects (Hanani et al. 2015).

Matras and Schiff (2005) conducted another variation study based on 10-50 minutes long recordings of 40 speakers aged 20-25 who did not pursue higher education, most from Mizrahi and peripheral backgrounds. They call the main vernacular in their sample "the working class (Mizrahi) vernacular" (parenthesis in the original label). Importantly, they acknowledge that some of the features of the "working class" vernaculars can be traced back to a substratum of North African Arabic (e.g., p. 185). The other "stylistic/register" points included in their analysis are

Formal (Normative) Hebrew, corresponding to a formal writing style, found mostly in essays; Educated Israeli Hebrew, used in formal contexts, such as lectures; and General Colloquial Hebrew, corresponding to the aforementioned General or Kóine vernacular). This seems to imply that all linguistic resources are on a continuum that should, in principle, be available to every speaker, or at least to all educated speakers. Is it truly the case?

As Yehudit Henshke noted in a lecture in 2021, speakers from different linguistic backgrounds within Israel don't necessarily share the same linguistic resources, and it is not always the case that educated Ashkenazi speakers commend the vernacular spoken by (educated or uneducated) Mizrahis from the peripheral parts of the country. In fact, as I show in Chapter 3, the phonological representations of native Hebrew speakers of different linguistic backgrounds may differ dramatically.

### **1.6 Third-wave sociolinguistics: identity construction in a segregated space**

“A British actor playing a cockney accent would produce a perfect imitation of a cockney accent, but an Israeli actor playing a commoner would settle for a caricature of a Mizrahi accent” - Yaron London, publicist and TV host

In an ethnography of two elementary schools in Northern California, Eckert (2008b) shows that producing Chicano vowels contrary to Anglo vowels does not simply index ethnicity, but also the speaker's place in the peer-based social order. A crucial component of the analysis is that using either variant is available to speakers regardless of ethnicity. The claim I will support in this section is that while an array of possible features may be available to speakers in less segregated environments, most of the features of speech that distinguish peripheral/Mizrahi from Center/Ashkenazi (and from some imagined, “middle”, native variant) are hardly controlled. In other words, it is in most cases challenging for speakers of either group to reliably produce the features of speech of the other. This, as the citation in the beginning of this section illustrates, is a longstanding tradition within representations of Mizrahis in the media.

Lefkowitz (2004) provides an analysis of language use in Israel as a process of identity construction. This is the lens through which every linguistic variable is examined in the current work. As reviewed above, there is a rich history of language patterning that suggests itself to such analyses. However, Lefkowitz reduces linguistic variation to a declarative level: since language use is seen above all in terms of identity formation and projection, significant differences, such as the difference between first language acquisition and second language learning, are described in similar terms. In fact, Lefkowitz explicitly objects to the separation between first and second language use, since “such a focus on individual ability and failure stems from a particular ideology of language that strictly opposes “mother tongue” to “second language”. This view of language erases the social context within which language is produced, thereby erasing also the social meaning situated speech generates (Lefkowitz (2004):132).

Since the process of language learning is a social-identity construction issue, Lefkowitz ties together Hebrew learning by Jewish native Hebrew speakers, Hebrew learning in the Ulpan (immersive language learning for Hebrew as a second language), his experiences of Hebrew learning in a Sunday school in Ohio, his efforts to speak Hebrew in Israel, Arabic acquisition by Palestinians, Arabic learning in the IDF intelligence units and English learning by Jews and Arabs in Israel. These experiences are put into a single framework of identity mediation, in which every language has its own social baggage: Hebrew means Zionism, innovation, and the state; Arabic is a “home” heritage language, symbolizing tradition, Palestinian nationalism and the political debate around it, and Jewish militarism; and English marks intellectual superiority and prestige.

As part of this perception, linguistic variables are not viewed as features of a natural language, but rather, as social cues that are being used consciously. For example, the use of pharyngeals is described as serving various goals: a manifestation of Ashkenazi-Zionist leadership, combined with a Middle Eastern vibe; a familiar attitude to fellow Mizrahis; and the integration of Palestinians in the Jewish-Israeli space:

*For the museum guide*, the pharyngeal forms semiticized his Hebrew. By using etymologically “correct” forms that are also shared by Arabic, the guide showed learnedness and sophistication at the same time that he symbolically placed Israel in its Middle Eastern context, laying claim to legitimacy for the Zionist presence in Palestine. *For the Mizrahi soldier*, the pharyngeal forms constituted an in-group language, through which he expressed and claimed solidarity with the Mizrahi taxi driver, whom he recognized as a member of his own identity group. *For the Palestinian businessman*, the pharyngeal forms arabized his Hebrew. By using elegant and prestigious Hebrew that simultaneously sounds like Arabic, the greengrocer claimed membership in Israeli society while simultaneously asserting a Palestinian identity (Lefkowitz 2004:222)

Based on sociolinguistic interviews, Lefkowitz determines that Palestinians use the pharyngeals when they speak Hebrew more often than the Jewish population (only 3/17 of the Jewish interviewees were Mizrahi). While this is likely to be true for the voiceless pharyngeal fricative [ħ], which tends to be produced categorically by speakers who do not merge it with [χ], the patterns of use of the voiced pharyngeal approximant [ʕ] are more varied in both groups, a fact that also seems to be closely tied with attitude and identity (e.g. Gafter (2016a); see Chapter 2). That is, like most of the Israeli society, Lefkowitz binds the pharyngeals together, thus ignoring significant differences in their patterning, and particularly the increased sensitivity of [ʕ] to matters of identity, compared with the linguistically categorical use of [ħ].

When discussing new forms of variation within the Hebrew speaking society, Lefkowitz describes a pitch contour that is initially associated with Mizrahi speakers, and then becomes a resource that spreads through society to project attitudes typically associated with Mizrahis, mostly agency and emotional involvement. The linguistic and social status of this variable resembles that

of Eckert's 2008b description of the use of the Chicano vowel system in California in two crucial points, that the status of the pharyngeals from the previous proposal did not share. First, the variable is available to all speakers of the community: in the Californian school case, since the children are exposed to varied speech and its structured social meaning from a young age; in the case of the Hebrew pitch contour, since pitch is a highly salient and more easily controlled feature of speech. The pharyngeals, in contrast, are harder for Ashkenazi speakers to produce, and [h] in particular is hard to produce *reliably* (i.e. not to accidentally produce it in cases that [χ] should be produced) when the etymologically correct phoneme is missing from the representation.<sup>12</sup> As with the choice of immigrants to reject the Yiddish penultimate stress pattern and adopt the Sepharadi pattern instead, salience and ease of production are likely to push speakers to adopt intonation more easily.

Second, the variables that speakers can cross ethnic boundaries with – vowels for English and intonation for Hebrew – are not the stereotyped variants of the dialect. In English, using a Chicano *intonation pattern* is likely to be heard as racist, due to the use of Mock Spanish in national television (Eckert 2008b). In Hebrew, the use of the pharyngeals raises similar concerns, again influenced by media representations.<sup>13</sup> Variants that spread across ethnic groups are likely to share this feature in particular - i.e. not being stereotyped - since alignment in speech features on an individual level usually takes place when the borrower perceives the speaker favorably (Lev-Ari (2016); Babel (2012); Babel et al. (2014)); adopting features that are used in the media to ridicule speakers is an unlikely strategy if the social goal is not hostile.

In the case of the intonation pattern, the analysis proposed by Lefkowitz (2004) seems more probable (see Chapter 2).

## 1.7 Conclusions and overview of the thesis

The discourse about variation in Modern Hebrew began when the language had very few native speakers. Accordingly, many forms of variation were discussed, stemming mostly from the first languages of European speakers, but also from the Sepharadi liturgical tradition – the sources of variation that were familiar to Zionist thinkers at the beginning of the 20th century. The goal of the discourse on variation at the time was to propose ideal productions that would set a standard for educators and learners. At the same time, the Zionist leadership and later, the state of Israel, put

---

<sup>12</sup> Since [h] and [χ] are merged in the Ashkenazi dialect, speakers often err when trying to speak with pharyngeals, such that they merge the categories again - but into the marked phone [h], as alluded to in the opening citation of this section. For evidence that merged speakers have different phonological representations of [h] depending on their linguistic background, see Chapter 3.

<sup>13</sup> A recent example (2015) is a video commercial for an apartment building in which you could finally “live with people like you”, in which a gentle Ashkenazi family is being harassed by two vulgar caricatures of Mizrahi men - in their previous apartment building. The Mizrahi men use the stereotyped pharyngeals. Every mainstream media outlet in the country has covered this story, since the usage of pharyngeals in this context was viewed as racist, e.g. <https://www.haaretz.co.il/gallery/media/1.2787911>, [https://www.calcalist.co.il/real\\_estate/articles/0,7340,L-3674790,00.html](https://www.calcalist.co.il/real_estate/articles/0,7340,L-3674790,00.html), <https://www.ynet.co.il/articles/0,7340,L-4775447,00.html>, <https://www.globes.co.il/news/article.aspx?did=1001084674>.

efforts into eradicating other languages from the landscape of Israel/Palestine. A recent example of these efforts is the “Basic Law: Israel as the Nation-State of the Jewish People” (also called “the Nationality law”, 2018). The law states that only Hebrew is the official language of the state, granting Arabic “a special status” (Arabic had been an official language until this point). Many appeals to the Supreme Court claimed that the law is unconstitutional, relying among other things on the fact that no state has ever rolled back the linguistic rights of its citizens, but the ruling was in favor of the state. In the court’s response to the appeal, the judges write:

A review of legislation, case law and literature shows, as the respondents (i.e. the state, S.B) point out, that even before the enactment of the Basic Law: Nationality, Israeli law gave priority to the Hebrew language over its Arab counterpart” (Nationality Law Judgment 9/7/2021, p.46, my translation)<sup>14</sup>

While Modern Hebrew spread as a native language, discussions on variation took a different form, distinguishing between variants that are associated with Ashkenazi vs. Mizrahi identities. The pharyngeals in particular were iconicized (in the sense of Irvine and Gal 2000) due to their transparency with respect to the orthographic system and the “Semitic vibe” on the one hand, and their marking of speakers from the Sepharadi and Yemenite traditions (liturgical as well as native language in most cases) on the other. The pharyngeals are at the intersection between the ideology of language planning (as in the back-and-forth between Jabotinsky and Yellin), and the (failure of the) melting pot, marking “non-integrated” individuals.

The mixture of language ideologies that has accompanied Modern Hebrew from before its emergence as a native language still affects the language and its speakers today. Language planning and prescriptivism, in the style of Jabotinsky, Yellin and the Battalion for the Defence of the Language is still found in government-funded commercials (e.g. by the ministry of health or road safety) and in public television and radio stations, all supervised by the Hebrew Academy. The Academy marks variation that is clearly related to the Arabic substratum as “erroneous”. For example, every couple of years the Academy’s Facebook page, its main outlet for direct interaction with the general public, insists that using the verb *limzog* ‘pour’ for serving solid food is wrong in Hebrew.<sup>15</sup> This usage of the verb was borrowed from the Arabic usage of *sakab*, as they acknowledge. The “authenticity” attitude of Yellin, i.e. the attitude that puts loyalty to the alleged historical dialect first, is still the primary justification used by the Academy when preferring one form over another, and this case is no exception: in the Biblical sources, *sam* ‘put’ is used for this purpose, and this is the source we should follow.

Other dimensions of variation do not meet the same vigorous objection. For example, broadcasters are no longer required to produce the pharyngeals and [r]. Influences from Aramaic are

---

<sup>14</sup>The full verdict is available here: [https://www.haaretz.co.il/embeds/pdf\\_upload/2021/20210708-175853.pdf](https://www.haaretz.co.il/embeds/pdf_upload/2021/20210708-175853.pdf)

<sup>15</sup><https://www.facebook.com/AcademyOfTheHebrewLanguage/photos/a.181203348616396/4326143017455721/?type=3&source=48>

accepted as well, despite being as ‘foreign’ as Arabic.<sup>16</sup> In addition to the aspiration for authenticity, then, Modern Hebrew prescriptive norms distance the language from more recent influences, and in particular those of Arabic.

While for some speakers of Hebrew it is clear that some variation arises from contact with Arabic, others rely on the Hebrew Academy’s view in seeing any divergence as an error. When taking up a prescriptive view, the same confound between low SES and Mizrahi/Peripheral speech, found in the education literature, seeps into the general public: if one strongly believes that there is a “correct” variant, those who use a different variant are wrong, due to some personal deficiency, lack of education etc. In discussions about “errors” on the Hebrew Academy’s facebook page, some participants highlight the prescriptive view while others claim that features of Arabic should be accepted into the language, since many of its native speakers use them naturally.

A particularly interesting example of the public discourse on variation revolved around a sketch in the most successful sketch show in the country (Eretz nehederet, lit ‘wonderful country’). The example is somewhat unusual, since participants in the discourse refer to phonological features, and not to syntactic and lexical variation that is usually at the center of attention. In the animated sketch, two characters portray traditional Hanuka pastries. They adopt a Mizrahi accent, while an enzyme, the authoritative figure that tries to educate them and restore order, is characterized by an Ashkenazi accent.<sup>17</sup> Although there is no difference in the character’s phonological inventory – in particular, none of them uses pharyngeal sounds) – audiences with increased sensitivity to the issue, e.g. a Mizrahi activists Facebook group and a feminist Facebook page, have referred to them as “sounding Mizrahi” and “sounding Ashkenazi”, respectively. The Mizrahi-sounding characters were portrayed as vulgar, ignorant, uncooperative with the authorities and finally, wearing scarfs and hats of an ultras football club whose main attribute in Israeli public discourse is overt animosity and racism against Arabs and Muslims – all attributes that are being consistently associated with Mizrahis (e.g., Ribke 2004).<sup>18</sup> It seems, then, that at least some listeners claim to being able to identify Mizrahi-sounding speakers based on phonological features other than the known markers. Others attribute the same forms of variation to SES or geography, as discussed above.

This discussion, and additional scattered characterizations of young speakers as “sounding Mizrahi” or “sounding Ashkenazi”, suggest that variation outside of the known markers of Mizrahi/Arab speech receives gradually growing attention. With the rise of commercial media, the internet, and notably “reality television”, more loanwords from North African varieties of Arabic find their way into colloquial Hebrew, joining the loanwords from Palestinian Arabic that have entered the language through the mostly-Ashkenazi Palmach, the pre-state Jewish army in Palestine

<sup>16</sup> e.g., <https://www.facebook.com/AcademyOfTheHebrewLanguage/posts/5617233171680026>

<sup>17</sup> This is in fact a recurring theme in the representations of Mizrahi and Ashkenazi characters, as noted by sociologist Ella Shohat (2001).

<sup>18</sup> References to the accents of characters in the sketch can be found in <https://www.facebook.com/591727857554934/posts/1713012635426445/> and <https://www.facebook.com/arsimbaim/photos/a.1528216187228720/1706521969398140/?type=1&theater>; some screenshots, including comments that disagree with the posting pages’ position (i.e. insist that the characters do not “sound like they’re Mizrahi or Ashkenazi”), are available in Appendix A.

(Henkin-Roitfarb 2011). There is a cultural revival associated with Mizrahis reviving their variants of Arabic, reflected in music and film (Shohat 2015).

The question of how, if at all, we can define a Mizrahi ethnolect/sociolect has yet to receive a clear answer, and it is equally not clear that such an answer is required. The literature does not determine how this linguistic variant is characterized, to what extent it is relevant to third and fourth generation of Mizrahi speakers, and whether or not it can be dissociated from ethnicity, socio-economic status and place of residence. A possible problem is that given that Mizrahi speech is more common in “the periphery”, and since “the periphery” is an ethnic and socioeconomic definition, speech becomes conceived of as the result of low SES, as if scarcity of resources somehow affects the speaker’s phonological inventory.

This assumption, I argued, should be turned on its head: Mizrahis of high SES are the ones that were under more pressure to adopt Ashkenazi features of speech. Mizrahis who were placed in peripheral, ethnically homogeneous environments retained and developed characteristics of speech that were influenced by their native languages, mainly Arabic. These peripheral locations were also underprivileged in terms of education and career opportunities, hence the correlation with poverty. On the other hand, Mizrahi speakers who lived in Ashkenazi-majority areas had the motivation to adopt Ashkenazi features of speech (or abandon characteristics that were stereotyped in an Ashkenazi-majority habitat), since in the context of the city Ashkenazi speech was more prestigious. An analysis along these lines was proposed in Henshke (2013c), who emphasizes the connection between characteristics of the Hebrew spoken in the periphery and those of Arabic spoken by Moroccan Jews. The local dominance of the “peripheral” dialect means that speakers of Russian or Amharic who had been put in this environment in the 1980-90s also acquired some of its features, in addition to potential influences of their L1, that will require future investigations.

Under Henshke’s analysis, both of the main assumptions of the current work are maintained: (i) that “peripheral Hebrew” is a variety that developed under the influence of an Arabic substratum; (ii) that second-generation Mizrahi speakers who were not part of one of these segregated communities did not maintain most of the Arabic features; and on the other hand, speakers of a variety of other languages were influenced by these features due to their place of residence and their local prestige within it. Henshke writes mainly about syntactic and lexical characteristics of the Mizrahi/Peripheral sociolect; a phonological investigation is due as well. Chapter 2 presents a preliminary investigation of the ethnic perception of phonetic and phonological variables in MH, using quantitative experimental measures. It joins the current chapter in attempting to provide an updated view of ethnically-conditioned variation in MH. The results uncover a new correlation between the rhythm and perceived ethnicity.

Part II (chapters 3 and 4) explores how known social markers of Modern Hebrew affect the online processing of speech, by speakers of various linguistic backgrounds. As a crossroads of languages, accents, language ideologies and attitudes, Israel-Palestine provides a unique site for exploring the relations between social and linguistic perception. The chapters addresses the following questions: Can early experience with multiple dialects yield measurably different phono-



logical representations? What role can our social expectations have on processing, i.e., can social identification change the manner and speed of lexical retrieval, given the listener's expectations?

Chapter 3 reports two experiments exploring the phonemic status of the voiceless pharyngeal fricative [ħ] across three groups of MH speakers: Mizrahi speakers who produce this marker; Ashkenazi speakers who do not, and whose exposure to it is limited; and most interestingly – second generation Mizrahi speakers, who do not produce [ħ] but were exposed to it since infancy. The experiment yielded both groups effects, such that second generation Mizrahi speakers mostly pattern with their parents who produce the pharyngeal; and speaker effects, such that expectations from a Mizrahi speaker were different than those from an Ashkenazi speaker. That is, both the linguistic background of the listener and the perceived identity of the speaker affected acceptance rates and reaction times in a lexical decision task.

Chapter 4 explores the question of speaker-oriented processing more specifically, using the Implicit Association Task paradigm (Greenwald et al. 1998). Two experiments were conducted, targeting stereotyped personae who use marked segments, one in a native context (“Old Mizrahi” vs. the “Radio Broadcaster” persona) and one in a non-native context (“Argentinian” vs. “Palestinian Arab”). In both cases, the personae share one marker (e.g. both “Old Mizrahi” and “Radio Broadcaster” speech includes [ħ]) and differ in the production of another variant (e.g. the “Radio broadcaster” produces some words with the diphthong [ej], while the “Old Mizrahi” produces the same words with a monophthong). The objective was to test whether the association between variants/markers could be mediated by the perceived identity of the speaker. The mediation effect was marginal; however, listeners in the Argentinian-Palestinian experiment were significantly more accurate with the Argentinian marker under the Argentinian guise, indicating that perceived identity can facilitate fine-tuning to specific acoustic characteristics. I conclude with a discussion of how the results from Part II, taken together, contribute new predictions to the dual-route model for auditory word processing presented in Cai et al. 2017.

## II

# Perception of Ethno-linguistic phonological variation in Modern Hebrew

In Chapter 1, I outlined some of the reasons for the general disregard of most forms of ethnically-conditioned variation studies on Modern Hebrew. I hypothesized that native Modern Hebrew listeners perceive ethnically-related phonetic and phonological variation in the language. I also discussed some of the reasons for why listeners might fail to notice such variation, or attribute it to other sources, such as SES and place of residence. The goal of the current chapter is to follow up on two threads that arise from this discussion:

1. Chapter 1 showed that the discourse on ethnically-conditioned variation in MH has narrowed down on a relatively small number of features from early on. In addition, it reviewed media reports from listeners arguing that they can hear that a speaker sounds like they are from a Mizrahi/Ashkenazi background, despite the lack of the stereotyped Mizrahi consonants in the speakers' speech. This raises the prediction that there is more variation in the perception of MH than previously discussed. The first question is therefore whether such effects are anecdotal, or more general. In other words, can listeners reliably rate speakers as Mizrahi or Ashkenazi without the presence of the known segmental markers, i.e. is there a general convergence of listeners' judgement for a given speaker?
2. Mimicking the variationist literature reviewed in Chapter 1, there is a wide range of perceptions in public discourse regarding the role of ethnicity in Israeli society. Do the view and social background of listeners affect their perception of variation? Particularly, do listeners who perceive ethnicity as insignificant also tend to be reluctant to rely on phonetic/phonological variation to identify speakers as sounding Mizrahi/Ashkenazi?

In Section 2.2, I present an ethnic identity classification task used to explore these hypothesis. Participants were requested to rate the likely ethnic identity of a speaker on the basis of a short recording (~5 seconds). The recorded voices, all native speakers of Hebrew, some bilinguals with an additional L1 (English, French, Yiddish, Polish or Arabic), were of Middle Eastern and European ethnic backgrounds, divided into two age groups (23-37 and 53-72). The older group provided a baseline for the task; since the older group includes Mizrahi speakers who produce the

segmental markers, and it has long been acknowledged that Hebrew speakers are very sensitive to these markers (see section 2.1), they should be rated as Mizrahi with the highest scores of certainty. Demographic and attitudinal data were also collected, in order to measure the effects of personal background and ideology on ratings in the task.

The chapter is organized as follows. In section 2.1, I review previous characterizations of ethnically-conditioned phonological and phonetic variation in Modern Hebrew, including but not limited to the stereotyped pharyngeals and /r/ (variants that have almost vanished from the inventory of young native speakers; see Chapter 1). Next, Section 2.2 presents the methods of the experiment, including an acoustic analysis of the stimuli and the demographic and attitude questions used to assess participants' background and attitudes. Section 2.2.2.1 includes an analysis and discussion of the inter-listener agreement regarding speakers, i.e., how reliably a speaker was classified as Mizrahi/Ashkenazi (question 1 above); Section 2.2.2.2 explores the possible contribution of various acoustic factors to ethnicity perception; and Section 2.2.2.3 includes the analysis and discussion of correlations between attitudes and performance in the classification task (question 2). Section 2.3 highlights new observations of the current study within the context of previous variation studies on Hebrew.

## **2.1 Previous studies of ethno-linguistic phonological variation in MH**

Most studies on phonological variation in Modern Hebrew focused on production rather than perception, and on segmental (primarily consonantal) rather than prosodic features. The current section reviews their findings, ranging from large-scale quantitative investigations to a small-scale perception task. While Chapter 1 focused on the narrative around variation in MH, the current section discussed the details in more depth. The sections are organized according to their methodology.

### *2.1.1 Quantitative production studies*

In terms of phonological variation, three consonants have been studied extensively using quantitative methods as markers of the salient ethnic divide between Mizrahi and Ashkenazi; these are [ʕ]~[Ø], [ħ]~[χ] and [r]~[ʁ], respectively.

The literature on variation in Modern Hebrew notes that the Mizrahi markers, i.e. the voiced pharyngeal fricative/approximant [ʕ], the voiceless pharyngeal fricative [ħ] and the apical trill/flap [r], are disappearing (Blanc 1968; Ben Tolila 1984; Davis 1984; Yaeger-Dror 1988; Lefkowitz 2004; see Table 2.1). This process dates back to the first generation of native speakers; by 1960, Blanc describes, 'there is a marked (but not total) leveling of communal differentiation, both in pronunciation and in other domains' (1968:240). As discussed in Chapter 1, the ongoing process of convergence is not a uniform phenomenon; many speakers in ethnically-homogeneous residencies retained their variety of Hebrew, influencing future generations. At the same time, it is

clear that by now most second and third generation Jewish Israelis of all ethnic backgrounds do not distinguish between [h] and the uvular voiceless fricative/trill [χ] in production, producing some variant of [χ] for both phonemes (Blanc 1968; Davis 1984; though see Ben Tolila 1984, a quantitative production study set in a rural Israeli town settled primarily by immigrants from Morocco, where pharyngealization rates were high for both adult and children).<sup>1</sup> They rarely or never produce [ʕ], defaulting to null instead.<sup>2</sup>

Unlike the other pharyngeal, [ʕ] was shown to be used stylistically: by some Mizrahi speakers, it is used more often in careful than in casual speech, implying that its usage is associated with some form of prestige (Gaftér 2016a). This is not surprising, considering that apart from being stereotyped sounds associated with Mizrahi identity, the pharyngeals are actually the prescriptive norm, and are supported by orthography (Yaeger-Dror 1988; Ben Tolila 2002); that is, the common pharyngeal-less dialect is actually the opaque one, orthographically speaking (see Table 2.1). In fact, two of the older Ashkenazi speakers in Gaftér's 2016a production study used [ʕ] when reading a word list, indicating that they too ascribe more prestige to this production under particular circumstances. It might sound somewhat surprising, however, that [ʕ] can be used stylistically, while there is no evidence that [h] may be used in the same way. This might be due to the fact that the alternative to producing [h] is producing [χ], thus merging the categories, while the alternative to producing [ʕ] is null or a glottal stop, which can more easily be perceived as a sort of reduction. Interestingly and in line with this hypothesis, the same pattern was found in Palestinian Arabic: while [h] is always categorically distinct from [χ] (indeed, only a unique historic trajectory could have yielded this merger), [ʕ] often undergoes lenition, and not only in productions of speakers who are in daily contact with Hebrew speakers (Horesh 2015).

Finally, most second generation speakers and onward produce a uvular/velar voiced approximant/fricative [ʁ] instead of [r] (Yaeger-Dror 1988; Cohen et al. 2019). The variant of /r/ used in speech is more ambiguous than the pharyngeals in determining ethnic identity (see below, in Section 2.1.3 and in Chapter 4).

Ashkenazi features of speech received less attention in the literature. That is because most features of old Ashkenazi speech have been adopted by succeeding generations and are consequently unmarked. One exception is a diphthong production of the historically long /e:/ vowel, sometimes called “*Tzere*”/ “*Tzejrej*” after the diacritic that represents it in the writing system (two horizontal dots under a letter).<sup>3</sup> Some older Ashkenazi speakers maintain a marginal phonemic distinction between the short and long vowel, producing the historically long vowel as [ej] and the short one

---

<sup>1</sup> The merger between [h] and [χ] will be discussed thoroughly in Chapter 3, where it is at the center of a discussion on the underlying representations of listeners to Merged vs. Pharyngeal dialects, i.e. listeners who were mainly exposed to the historically Ashkenazi or Mizrahi consonant inventory, respectively.

<sup>2</sup> The same reduction to null happens with most occurrences of the glottal stop and glottal fricative [h]. Contrary to variation in the production of [ʕ], variation in [h] deletion rates has never been associated with a particular ethnic identity to my knowledge. It was, however, studied both quantitatively in an acoustic analysis of multiple speakers and in perception in Gaftér 2014; see the bottom of section 2.1.2.

<sup>3</sup> There is a debate regarding the original production of this vowel. However, it will suffice us to note that it is orthographically marked with an additional vocalic letter compared with the “short” /e/ in Tiberian Hebrew.

Hebrew letter	Historically Mizrahi	Historically Ashkenazi (Common)
aleph (א)	ʔ	ʔ~∅
ʾayin (ע)	ʕ	
kaf (כ)	χ	χ
heit (ח)	ħ	
resh (ר)	r	ʁ

**Table 2.1: Markers of Mizrahi speech and their corresponding letters. The historically Ashkenazi dialect, which is the more prestigious one in most social contexts, is the one that includes mergers (ħ → χ, ʔ/ʕ → ∅). Most young native Hebrew speakers use the common dialect, while only few young (Mizrahi) speakers use the marked [ħ], almost none use [r]. The usage of [ʕ] is stylistically conditioned and sporadic (Gafter 2016a). All of these variables have previously been acoustically analyzed and studied using quantitative measures.**

as [e].<sup>4</sup>

While some young speakers retain [ej] in a few common words (e.g., *tejfa* ‘nine’, *tej* ‘tea’), its status as a full-blown native phoneme has been gradually diminishing since at least the 1950s, a process that seems to have proceeded faster in the city than in rural Ashkenazi communities, like the Kibbutz (Matras and Schiff 2005; Neuman 2012). As with the Mizrahi markers, most second and third generation speakers do not systematically distinguish /e/ and /ej/. In a production study that included speakers from Mizrahi and Ashkenazi backgrounds, Matras and Schiff 2005 found that out of a total of 1028 tokens containing the etymological diphthong, only 3-29% (depending on the speaker) were produced as such. Unlike the Mizrahi markers, the diphthong is a gradient variable: it is affected by lexical factors, and its production is less foreign since it is mediated by the existence of /j/ as a glide in the language. I speculate that the greater salience of Mizrahi markers is to some extent also derived by this: [tej] ‘tea’ can be parsed as a valid CVC structure even if [ej] is not interpreted as a diphthong; words that include a /j/ coda exist in all varieties of Modern Hebrew, e.g. /jamaj/ ‘sailor’. On the other hand, [ham] ‘hot’ cannot be parsed within the phonemic repertoire of Hebrew in any other way but as the Mizrahi production of this phoneme. In other words, every element in [tej] exists in the phonological inventory of Hebrew, but not every element of [ham].

Despite its declining use and gradient nature, it seems early to eulogize the [e]~[ej] distinction as a possible sociolinguistic indicator. The prevalence of borrowed words from English marks a new era for the status of the diphthong [ej], that may or may not be produced in English loanwords such as o.k, e-mail, chaser, HD television, Facebook. New loanwords are often related to technology, and are often incorporated faster by the younger generation. In fact, the majority of diphthongs in Matras and Schiff (2005) were found in loanwords. The emergence of new words that include the diphthong [ej] raise the nearly forgotten [e]~[ej] distinction back to the surface, and introduces a new layer of sociolinguistic meaning to the variable: connection with

<sup>4</sup> Apart from this alleged remnant to vowel length in some variants of the language, Hebrew does not have length distinctions.

the English-speaking world. According to Matras and Schiff (2005), some loanwords are also produced without the diphthong, suggesting that this variable as well may also affect social perception.

### 2.1.2 *Qualitative studies and impressionistic observations*

The remaining observations regarding ethnicity-related variation in MH, outside of the features described in Section 2.1.1, are prosodic. The term “prosody” (also “supra-segmental features”) is used as a cover term for features of speech that are not segmental: tone, stress, length, and intonation. It has long been acknowledged that languages differ in their rhythm (Pike 1945), and that differences in rhythm are highly perceptible. For example, Ramus and Mehler (1999) showed that even when the phonemic and intonation cues are dramatically degraded ([s] was re-synthesized instead of every consonant/cluster, and [a] instead of every vowel), syllabic rhythm and syllable structure can suffice for listeners to distinguish between languages from different rhythm classes.

Vernaculars of the same language can also vary in their rhythm to a degree that is perceived by speakers. In English, discrimination between American and Australian varieties was significantly above chance when only rhythmic cues were available (Vicenik and Sundara 2013). In Arabic, naïve native speakers from various regions were able to discriminate between Middle Eastern and North African dialects 98% of the time (Barkat et al. 1999), and their classification into different dialects within each of these broad categories (e.g., Tunisian v. Moroccan) was shown by Hamdi-Sultan et al. (2004) to correlate with fine-grained distinctions in the proportion of consonants vs. vowels. In addition to rhythm, pitch may also contribute to social perception. De Mareüil and Vieru-Dimulescu (2006) manipulated both prosodic factors of Italian and Spanish natural utterances, and found that native listeners of these languages classified the utterances in line with prosody. The recorded sentences included a limited repertoire of segmental variation between Spanish and Italian, and under these circumstances, syllable duration and pitch contour played a more crucial role in discriminating between the languages than segmental variation. Swiss German speakers were also able to discriminate between three out of four dialects of their language using pitch contour and rhythm (i.e., no segmental cues), with particularly high success rates when the dialects exhibited unique prosodic phenomena, such as a significantly slower speech rate or a large pitch range (Leemann and Siebenhaar 2008).<sup>5</sup> <sup>6</sup>

Subtle prosodic variation was shown to play a role in dialect classification in the US, even with very short samples of speech; in fact, in a series of four experiments, Purnell et al. (1999)

---

<sup>5</sup> De Mareüil and Vieru-Dimulescu 2006 proposed that speech rate – and more particularly, lengthening of stressed syllables – is what allows a large pitch range. The relation between these prosodic properties might be similar in Hebrew, in which the HLH contour is correlated with rhythm; see Section 2.2.1.

<sup>6</sup> However, in a follow-up experiment which forced speakers to choose between segmental and prosodic features for classification, since it involved swapping the rhythm of one recording with the other (Leemann et al. 2018), speakers have for the most part neglected the prosodic dimension and relied almost exclusively on segmental differences, unlike in the Italian/Spanish discrimination task in De Mareüil and Vieru-Dimulescu 2006. One possible reason is that segmental variation in the latter was minimal, while rhythmic variation was particularly robust.

showed that American listeners were well above chance in classifying speakers into the African American vernacular (AAVE), Standard American English (SAE) and Chicano English (ChE), even when only a recording of the word “hello” was presented to them. As in the current study, the authors were particularly interested in phonetic features that were not known markers of any of the three dialects. Their main objective was to test whether speakers can be discriminated against in housing based on a very short exposure to a voice, which is often the case when calling about an advertised apartment. Indeed, they found a strong correlation between the used dialect, neighborhood and appointments scheduled: apartment renters in majority-white neighborhoods tended to reject speakers of AAVE and ChE (based on hundreds of phone calls made by the tri-dialectal last author of the study). In another experiment, Purnell et al. 1999 confirmed that young Caucasian listeners were, as mentioned above, quite good at identifying the speaker’s ethnicity based on the word “hello” alone. However, the acoustic data were more difficult to interpret. Out of 28 measures, four were found significant, and two were easier to interpret. One was the peak pitch within a syllable, which is perceived as a correlate of stress in English: the AAVE “hello” peaks earlier and could therefore be seen as more trochee-like. Another was the length of the first syllable, also longer on average in AAVE recordings. The acoustic data could not account for confusability of AAVE with SAE, which probably means that there were other minute differences of voice quality between the guises that are just not easily detectable by the acoustic analysis.

In sum, research on both language and accent classification from prosodic cues confirms that rhythm is the most reliable prosodic feature – or at least, the easiest to measure – with pitch accent usually playing a more modest role; and both seem to be less reliable/salient social cues for speakers than segmental cues in the general case. The American English data hints that they can be a rich source for ethnic classification, but at the same time, that they are challenging to study.

Perhaps unsurprisingly then, the dimension of prosody in variation studies is often overshadowed by segmental variation, as is the case in Modern Hebrew. Nevertheless, the public’s attention to prosody seems to be on the rise, with an increased use of terms such as “the Mizrahi tune” or “Mizrahi intonation”. In an article from 2002 about the “Ars” (a pejorative term used against Mizrahi men), an interviewee describes: “...[the] Ars is a person who *elongates* words. He will say every word in the longest way possible - *ma ko-re it-xa* (‘what’s up with you?’), very slowly” (my emphasis).<sup>7</sup> The impression that Mizrahi speakers “elongate” words echoes Bolozky (2002), who claims (based on personal observation; the work does not include acoustic measurements) that Yemenite and Ethiopian Hebrew speakers tend to elongate the pre-tonic syllable. Further, Bolozky notes that this characteristic makes Mizrahi Hebrew closer to syllable-timed languages such as Spanish. At the other extreme, some Ashkenazi speakers produce Hebrew as more close to the stress-timed class – perhaps under the influence of Yiddish and German, and possibly due to the recent influence of English.

Pre-tonic lengthening is also characteristic of young speakers of French from an area in which

---

<sup>7</sup>The full article (Hebrew) is available at: <https://www.haaretz.co.il/misc/1.822301> (last accessed on 1/3/2021).

most speakers are second and third generation to immigration from north Africa (Fagyal 2003). In addition to the verification that such data provides as to the origin of pre-tonic lengthening, the fact that the same feature received attention in the literature on variation in French suggests that it is salient enough for listeners to track.<sup>8</sup>

Pitch has also been claimed to be used differently by Mizrahi and Ashkenazi speakers, although less categorically. The distribution of one contour in particular was proposed by Lefkowitz (2004) to be ethnically-conditioned: the “rise fall rise tune”, or \*HLH%, which he describes as a three-step tune: “(a) a high pitch-accent aligned with a stressed syllable, (b) a fall to a relative low pitch close to the end of the phrase, (c) a sustained high pitch over a lengthened phrase-final syllable”. Hebrew speakers tend to use high nuclear and rising tones regularly, regardless of ethnicity (e.g. Ozerov 2013). It is therefore necessary to specify the singularity of this contour, beyond Lefkowitz’s description. Based on my small sample of acoustic data, an additional crucial feature of the contour is that the phrase-final boundary tone must not peak higher than the nuclear high pitch at the beginning of the phrase. This progression of \*HLH% is demonstrated in Figure 2.1, from a speech sample of a young woman recorded for the current study.

Lefkowitz (2004) notes that in his sample, Mizrahi speakers use the contour regularly (once every couple of sentences), and do not require a particular pragmatic context in order to use it. Ashkenazi speakers, on the other hand, used it more rarely and in specific contexts, for the most part when they wanted to emphasize their involvement or agency. This, according to his analysis, indicates that Ashkenazi speakers tend to perceive Mizrahis as more emotionally involved and more agentive, and the borrowed pitch accent therefore conveys attitude. This interpretation is compatible with the way Mizrahi characters are construed in media representations in the past decade, as assertive and authentic (e.g., Mehager 2018; Levon and Gafter 2019).

Another observation that has not been reported in the literature so far is that some young speakers, particularly in the Tel Aviv area (i.e. the Urban center), adopt a speech style that includes a high proportion of rising boundary tones - not unlike the speech style dubbed ‘uptalk’ in the American English discourse. Ever since it drew media attention (Gorman 1993), Uptalk has stereotypically been associated with young women, which in most cases studied in the US indeed tend to use it more (Sando 2009; at least, this is the case in spontaneous speech, see Prechtel and Clopper 2016), and with Southern California (Ritchart and Arvaniti 2014). Nonetheless, it has been shown to be prevalent in many English dialects, including Australian, Southern England (Fletcher et al. 2005), and various American dialects (e.g., the American Mid-west: Prechtel and Clopper 2016, Canada: Shokeir 2008). Similarly, Modern Hebrew speakers often use rising boundary tones (e.g., Ozerov 2013, 2019), but this particular usage — in which the general Modern Hebrew tendency of rushing at the beginning of the sentence and lengthening at the end is taken

---

<sup>8</sup> To me, it seems plausible that pre-tonic lengthening in Hebrew and French shares a source (according to Fagyal, West Arabic dialects). Bolozky (2002) attributes this feature to Ethiopian speakers (most of whom have Amharic as a heritage language); and to the Yemenite heritage population. Both possibilities require a more in-depth quantitative investigations into the current productions of speakers of these various linguistic backgrounds.



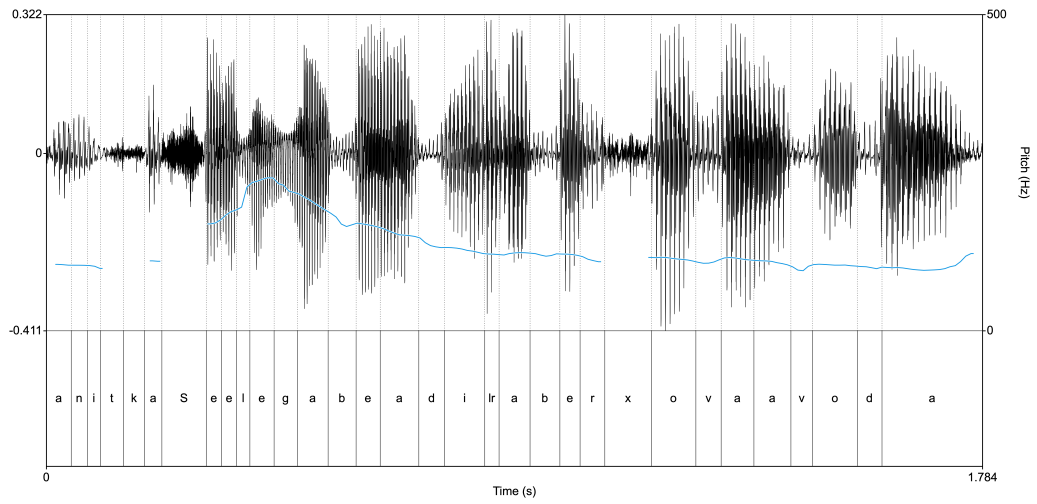


Figure 2.1: An example of the \*HLH% contour

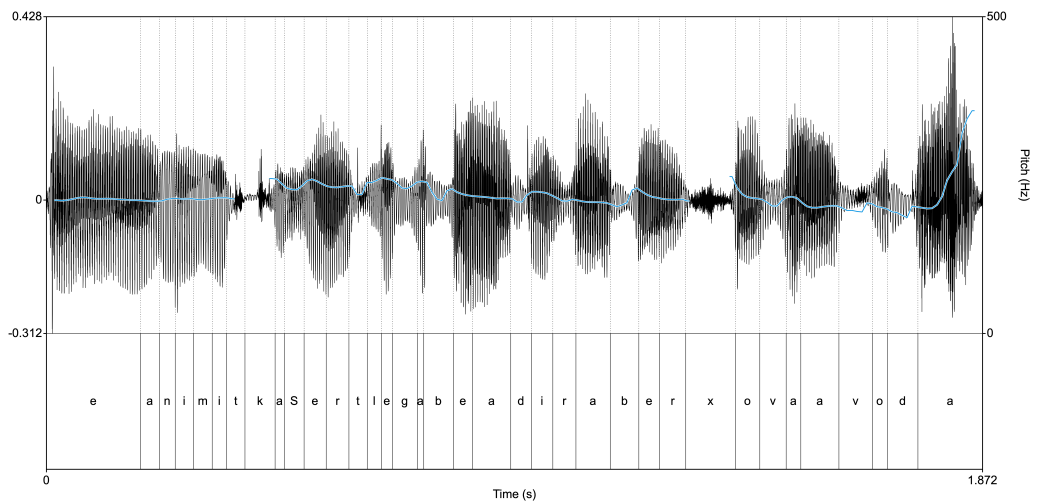


Figure 2.2: An example of the \*LH% rising contour

to an extreme – seems to be on the rise.<sup>9</sup> An example of this contour is available in 2.2.

### 2.1.3 Perception studies

The three Mizrahi consonants [ʕ], [ħ] and [r] are to my knowledge the only sounds in Modern Hebrew that have been studied as markers of ethnicity perception.<sup>10</sup>

Even before considering the perceptual data, there is ample evidence that the presence of the three marked Mizrahi consonants has a robust effect on perceived ethnicity. First, there is a common term for describing this variable – *ledaber be-heit ve-'ajin*, lit. ‘to speak with Heit and Ayin’; Heit and Ayin are the names of the letters that correspond to the pharyngeal voiceless fricative and the pharyngeal voiced approximant in the Mizrahi inventory. The rhotic variant that Mizrahi speakers use also has a unique name – “resh mitgalgelet”, lit. ‘rolling R’ – but this variable is more ambiguous with respect to ethnic identity, because it is used by various groups of speakers: Mizrahi, but also native speakers of Spanish, as well as speakers who do not produce it normally but only in official settings, most notably old school radio broadcasters (see Chapter 4). In addition, not all Mizrahi speakers who produce the pharyngeals use this variant (e.g. Yaeger-Dror 1988).

One of the key characteristics of a socially meaningful variable is that speakers can refer to it explicitly (Labov 1972) — and Hebrew speakers do so consistently with the pharyngeals, and to a lesser extent with [r]. Additionally, these variables seem to be highly salient – so much so that listeners use the label *ledaber be-heit ve-'ajin* to describe even speakers that produce a vanishingly small rate of pharyngeals. For example, a participant in a reality TV show who produced [ʕ] less than 2% and *never produced [ħ]* was commented on as “a person who speaks with Heit and Ayin” (Gaftner 2016b).

Getting to somewhat more direct evidence, two small-scale perception studies have been conducted that tested the evaluations of Mizrahi speech by native listeners. The first experiment used a matched-guise design in which high-school students from a majority-Ashkenazi Jewish school in Tel Aviv, and from a majority-Christian Palestinian-Israeli school in Jaffa, rated extracts of speech on personal traits, such as friendliness, trustworthiness and intelligence (Lambert et al. 1965). The speakers in the recordings were two bilingual high-school students: a native speaker of Arabic,

---

<sup>9</sup>The tendency to rush at the beginning of the sentence and slow down towards the end is also not unique to Hebrew (Izre'el and Silber-Varod 2009).

<sup>10</sup>Gaftner 2014 showed that the production of a glottal stop [ʔ] instead of the fricative [ħ] was negatively evaluated by MH listeners. Although this variable has not been directly associated with ethnicity, I point it out since it is one of the few recent studies of perception of variation in Hebrew. In a matched-guise design, listeners gave judgments of the same voices with one of three pronunciations, created by splicing: [ʔ], [ħ] or [Ø] (Gaftner 2014). Interestingly, in the interviews conducted and transcribed in Gaftner 2014, /h/ was deleted in 92% of its potential occurrences, making [Ø] the least marked variant; and indeed, speakers were not “punished” in terms of social judgment for not pronouncing [ħ]. However, the same voice was judged to be less nice and sophisticated when [ʔ] was produced instead of [ħ]. The less-favorable pronunciation was not overtly associated by listeners with a particular ethnic identity; then again, it might not be pleasant for a participant to relate ethnic identity and unpleasant social characteristics to the same speaker, since it would suggest they are biased. For this reason, I used different tasks for social and ethnicity judgements.

who was judged by the researchers to sound Yemenite (which falls under the “Mizrahi” umbrella term) when speaking Hebrew; and a native speaker of Ashkenazi Hebrew who was judged to sound native when speaking Arabic (allegedly; it is not specified who had made the judgement). Participants were unaware that they heard the same voices twice, and rated a total of 4 recordings (+ an unrelated filler voice) of the same story, a couple of minutes long. It should be noted that the authors do not state which “Yemenite” variants the native Arabic speaker had used; however, since native Arabic speakers often produce pharyngeals and [r] when speaking Hebrew, as they do in their native language (Horesh 2015), it is highly probable that the “Yemenite” guise included all three markers. According to the results, the Ashkenazi guise was rated as more good-mannered, confident, clean, wealthy, reliable, intelligent, ambitious and successful; the Yemenite guise was rated as more good-hearted, honest and with a sense of humor. These stereotypes are closely linked with stereotypes about Mizrahi people, as they are reflected in media representations of the time (see Shohat 2001 and Chapter 1 for discussion). The findings of this experiment are not direct evaluations of the speaker’s ethnic background; they are evaluations of the speaker’s traits. However, the underlying assumption of the experiment is that listeners can identify the ethnic background of the speaker themselves, and the results support this assumption. The fact that the evaluations of Yemenite/Mizrahi speech converge with observations regarding stereotypization of these identities in society is another indication that the listeners correctly identified the intended guise. Experiment 2 (Chapter 3) provides direct evidence that listeners of all ethnic backgrounds not only register the pharyngeal [ħ], but can also reliably distinguish it from [χ].

The second perception study, Devens (1981), is to date the only study that had similar goals to the current one: test whether Hebrew listeners can identify the ethnicity of a speaker, including when the known segmental markers [ʕ] and [ħ] are absent. In her design, 10 participants both recorded and listened to recordings of each other (i.e. each participant classified the other nine), in two settings: a word list which was set as to not include [ʕ] and [ħ], and a free conversation in which these variants occurred (though the author doesn’t state how often, or whether they occurred in the speech of *all* participants). In the free conversation setting, additional factors other than phonology could have interfered, including the topic of the conversation and lexical/syntactic choices. While the word list did not include [ħ] and [ʕ], /r/ was included in some word(s) and the author notes that it was produced by some Mizrahi speakers as the marked trill (it is not clear how many of them). The full set of data, simplified such that Yemeni, Iraqi, Kurdish, Moroccan, Persian as well as “Mizrahi” were all considered Mizrahi, and Romanian, French, English, and “European” were considered Ashkenazi, is available in table 2.2).<sup>11</sup> Devens argues that subjects performed poorly, based on the fact that they did not always converge to the “correct” ethnic background of the speaker. It seems to me, though, that participants generally agreed on the general ethnic backgrounds of speakers:

- In the word list evaluation, participants’ levels of agreement regarding a particular record-

---

<sup>11</sup> It is interesting to note that for “Ashkenazi” responses, the umbrella term was more common than naming a specific ethnicity; the opposite was true with “Mizrahi” responses, that tended to refer to specific ethnic backgrounds, e.g. Iraqi.

Speaker	Word list		Conversation	
	Mizrahi	Ashkenazi	Mizrahi	Ashkenazi
A	4	5	8	0
B	6	2	8	0
C	6	2	7	0
D	2	6	7	0
E	8	0	8	0
F	4	2	5	1
G	1	7	3	4
H	2	5	1	6
I	1	7	1	6
J	3	5	0	7

**Table 2.2: Number of listeners evaluating each speaker as Mizrahi or Ashkenazi in Devens 1981’s perception experiment, based on two different recordings: reading a word list or natural conversation.**

ing seem to be above chance (although there are not enough evaluation for a quantitative analysis): speakers B-F got mostly Mizrahi evaluations, speakers G-J got mostly Ashkenazi evaluations, and the recording from speaker A was exceptional in the sense that its evaluations are closely tied. It is not clear how many of the recordings included the marked /r/ variant. Yet, it is quite remarkable that without the pharyngeals, and in the absence of many prosodic cues that are only available in full phrases, some speakers have still been evaluated as sounding Mizrahi/Ashkenazi.

- In the recorded conversation, results were more robust and more leaning towards “Mizrahi” responses. This is probably due to the presence of the pharyngeals, but it can also be due to other properties not related to the phonetic/phonological level (e.g. lexical or syntactic characteristics, the topic of the conversation), since there was no control of these factors.

Even if it were correct that these particular speakers could not be associated with a particular ethnicity in the word list condition, as Devens argues, it is by no means enough in order to conclude that the pharyngeals alone determine ethnicity perception. First, some prosodic variables that may influence ethnicity perception are at the level of the phrase, such that they could not arise in a word list reading setting (see Section 2.1.2). Second, reading usually induces careful speech, that is more conscious and performative, and therefore may reflect attitudes and language ideologies more than other modes of speech (Schilling-Estes 2008; Gafter 2016a).

My reading of Devens (1981) suggests that listeners in the 1980s were sensitive to linguistic variation above and beyond the known markers. First, all the listeners in the study have at least *tried* to guess the ethnicity of each speaker, and admitted to not have an idea very few times (5 times out of a total of over 160 ratings, omitted from Table 2.2). These few times were also the cases that were more ambiguous in general, that is, that received mixed responses from listeners as Ashkenazi/Mizrahi. Second, some speakers were rated as more Mizrahi/Ashkenazi even in the

word list task. This is the more crucial point: considering “accuracy” alone in this case doesn’t seem to be the best way to answer the main question that arises regarding the perception of variation, which was whether Hebrew listeners can extract perceived ethnicity from speech. I suggest that a better measure would be inter-listener agreement. That is because speech is shaped by the speaker’s linguistic background, that includes not only their home, but also school, neighborhood etc., and therefore an Ashkenazi speaker in a majority Mizrahi school might align with their peers, and vice versa.

In sum, the perceptual data corroborates evidence from production and speakers’ overt reflections on the pharyngeals. There is less conclusive evidence regarding /r/, especially because it is not clear in either study whether the guises that were construed or judged as Mizrahi used the apical trill or not. Other ethnically-conditioned variants have not been examined systematically, but as reviewed above, based on their distribution in the different ethnic groups, they may indicate ethnic identity to the attentive listener.

Table 2.3 summarizes the phonological features discussed in the literature as possible indicators of ethnicity. Of those, only the three variables labelled as markers of “mostly old Mizrahi” speakers have been extensively studied, with many production studies, acoustic analyses, discourse analysis, as well as a few perception studies, all confirming the relations between these variables and perceived Mizrahi identity. Some production data regarding the [e]~[ej] distinction were collected in Matras and Schiff (2005), which correlated with ethnicity to some extent in their sample, but there is no evidence of perceived ethnicity for this variable.

Social group	Marker	Unmarked alternative	Reference
Mizrahi (older group)	[h]~[χ] distinction	merged [χ]	Blanc 1968; Ben Tolila 1984; Davis 1984; Devens 1981
	[ʕ] Alveolar trill/tap	[Ø] Dorsal approximant/fricative	“” Yaeger-Dror 1988; Matras and Schiff 2005; Cohen et al. 2019
Mizrahi	Pre-tonic lengthening *HLH% contour		Bolozky 2002 Lefkowitz 2004
Ashkenazi (older group)	[e]~[ej] distinction	Gradient merger in native words; often preserved in English borrowings	Neuman 2012; Matras and Schiff 2005
Young Ashkenazi	*LH% contour		My observation

**Table 2.3: Summary of proposed socially meaningful phonetic variables in Modern Hebrew**

Other variables were not studied to the same extent. Lefkowitz (2004) provides an interesting

discussion of the choice of pitch accent as a means of attitude projection; however, he offers no direct insight on perception, excluding his own. Bolozky (2002) hypothesizes that pre-tonic lengthening, and perhaps rhythm more generally, is ethnically conditioned, but does not provide supporting acoustic data. My observation regarding the increasing usage of rising contours among young speakers is not currently backed by quantitative data either. In the experiment presented below, the recorded speakers varied with respect to these features.

## **2.2 Experiment 1: Ethnic classification in ~5 seconds**

The first experiment was a classification task, in which subjects rated based on a short recording how certain they were that a person was Mizrahi/Ashkenazi, when these identity labels serve as the opposite ends of the scale, the mid-point being “I can’t tell”.

The main goal of the study was to test whether listeners would be able to reliably classify according to perceived ethnicity short samples of speech, which do not contain any of the known markers associated with ethnic identity in Modern Hebrew. This amounts to asking the question: will listeners of Modern Hebrew systematically classify the same recordings as sounding Mizrahi/Ashkenazi? A positive answer would strengthen our hypothesis that there are cues within the acoustic signal that affect perceived ethnic identity, beyond the known markers.

A secondary goal was to test whether the social attitudes and linguistic backgrounds of listeners play a role in their classifications. Particularly, I compared listeners based on their social background, attitude towards the Israeli ethnic cleavage and willingness to acknowledge phonological variation.

Finally, the experiment is also a first exploration of the acoustic cues that might be relevant for perceived ethnicity outside the known segmental markers. Several proposals have been made in the literature of acoustic features that characterize Mizrahi/Ashkenazi speech (see Section 2.1), but they have not been tested perceptually. In the acoustic analysis, I explore whether these and newly proposed acoustic indicators are also likely to be used by listeners, by correlating the ratings that a given recording received with the relevant acoustic features within it. These correlations can then serve as a new starting point for future investigations of prosodic variation in Hebrew.

### **2.2.1 Methods**

**Participants:** One hundred and ten native Hebrew speakers volunteered to participate in the experiment online. Five of them reported in the post-task comments section that they had not read the instructions well in the beginning, and warned that their results for the first trials might be affected. They were excluded from the analysis, leaving 105 participants. Their demographic data are summarized in Table 2.4. 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.<sup>12</sup>

**Stimuli:** Twenty-eight Native Hebrew speakers were recorded producing the utterance below. Hebrew has gender marking on the verb (in parenthesis below), which was produced or not depending on the speaker’s gender. Segmental markers are in boldface.

halo? jalom, ani mitkafer-(et) legabe**j** ha-dira be-rehov ha-ʕavoda.

hello? hi, I call-(fm.) about Det-apartment in-street Det-Avoda.

‘Hello? hi, I’m calling about the apartment in Ha’Avoda street.’

Speakers were of varied linguistic and ethnic backgrounds, more or less evenly divided across age and gender identity, which included only cisgender men and women (see Table 2.5). Of the fourteen speakers in the older group (53-72), six used at least one the ethnically-marked segments [ħ], [ʕ] and [r]. In the younger speakers group (24-33; 14 speakers), none used the marked segments. Speakers were instructed to memorize and then produce the sentence without looking at the screen on which it was presented. Three productions of each speaker were recorded and the most natural-sounding one was selected for use in the experiment (mean length of recording = 5 seconds).

**Acoustic analysis of the stimuli:** Speech samples were fully segmented on PRAAT (Boersma 2006), and the variables listed in Table 2.3 were submitted to an acoustic analysis. The variables listed in Table 2.3 were submitted to an acoustic analysis. In what follows, I detail how each measure was calculated. Table 2.6 summarizes the relevant dimensions of variation in the recordings, by ethnicity and age group.

The segmental markers [ħ], [ʕ] and [r] were identified by ear, since their Ashkenazi counterparts are significantly different acoustically. Perhaps surprisingly, the only speaker who used [ʕ] was of European background, and was a native speaker of French, in addition to Hebrew. As discussed in 2.1.1, [ʕ] can be used stylistically, e.g. be reduced in various conversational settings and reemerge in others; in fact, two Ashkenazi speakers in Gafter’s 2016a word list production task also produced [ʕ], though rarely. The other Mizrahi markers were distributed according to expectation, with 2 out of 7 older Mizrahi speakers producing [ħ] and [r], and 3 additional older Mizrahi speakers producing [r], for a total of 6 speakers producing the Mizrahi markers.

<sup>12</sup>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.

Ethnicity	Age (sd)	Total N (of whom women)
Ashkenazi	36 (9.8)	32 (18)
Mixed	34 (6.9)	23 (19)
Mizrahi	33 (8.3)	44 (40)
Other	37 (11.6)	6 (5)

**Table 2.4: Self-reported demographic data of participants in Experiment 1**

Label	Age	Gender	Environment	Languages	Parents languages
21	72	M	Ashkenazi (kibbutz)	Hebrew	Hebrew, Yiddish
22	68	W	Ashkenazi (city)	Hebrew, Yiddish	Hebrew, Yiddish, Polish
2	68	W	Ashkenazi (city)	Hebrew, Polish	Hebrew, Polish
8	60	M	Ashkenazi (city)	Hebrew	Hebrew, Yiddish
7	60	W	Ashkenazi (city)	Hebrew	Hebrew
9	56	M	Ashkenazi (city)	Hebrew	Hebrew, Czech
10	54	W	Ashkenazi (city)	Hebrew, French	Hebrew, French
18	29	W	Mixed (city)	Hebrew	Hebrew
25	28	M	Ashkenazi (city)	Hebrew	Hebrew
16	28	M	Ashkenazi (city)	Hebrew, English	Hebrew, English
17	26	W	Ashkenazi (city)	Hebrew, English	Hebrew, English
24	25	M	Mixed (city)	Hebrew, English	Hebrew
13	25	M	Ashkenazi (settlement)	Hebrew	Russian
12	24	W	Ashkenazi (city)	Hebrew	Hebrew
1	70	W	Mizrahi (city)	Hebrew	Yemeni Arabic
4	68	M	Mizrahi (city)	Hebrew, Iraqi Arabic	Hebrew, Iraqi Arabic
3	66	W	Mizrahi (city)	Hebrew	Yemeni Arabic
14	56	M	Mizrahi (Moshav)	Hebrew	Western Arabic, French
6	56	M	Mizrahi (city)	Hebrew, Arabic	Hebrew, Arabic
15	53	W	Mizrahi (city)	Hebrew	Hebrew, Western Arabic
5	48	M	Mizrahi (development town)	Hebrew	Western Arabic
27	33	M	Mizrahi (development town)	Hebrew	Western Arabic
19	30	W	Mizrahi (city/settlement)	Hebrew	Farsi
26	28	W	Mizrahi (city)	Hebrew	Hebrew, Western Arabic
28	28	M	Mizrahi (city)	Hebrew	Hebrew, Arabic, Bulgarian
23	28	W	Mizrahi (city)	Hebrew	Hebrew
11	26	W	Ashkenazi (Moshav)	Hebrew	Hebrew with pharyngeals
20	25	M	Mizrahi (city)	Hebrew	Hebrew, Yiddish

**Table 2.5: Personal background of recorded speakers, ordered by ethnic background and from oldest to youngest. Environment = the majority population in the area in which the speaker grew up (Moshav and Kibbutz are rural settlements, the latter of which used to be communal and hence more segregated).**

Age group	Mizrahi background		Ashkenazi background	
	24-30	53-72	24-30	53-72
[h]	0	2	0	0
[ʕ]	0	0	0	1
[r]	0	5	0	0
Pre-tonic lengthening	4	3	0	1
*HLH%	5	3	2	0
[ej]	0	0	0	0
*LH%	1	0	4	1

**Table 2.6: Number of speakers (out of 7 in each cell) that exhibit possible ethnicity markers, by speaker's age and ethnic background.**



None of the speakers produced the diphthong [ej]. This does not necessarily mean that the speakers in the sample do not make the [e]~[ej] distinction elsewhere. Rather, it seems more likely that the item with the potential environment for [ej], [legabej] ‘about’, was not revealing, because it is a function word, a fact that tends to make words prone to reduction.

The proposed prosodic indicators of Mizrahi speech were measured as follows. Pre-tonic lengthening (Bolozky 2002) was measured as the proportion between vowel duration in the word “dira” (V1/V2).<sup>13</sup> For purposes of presentation, the cutoff point above which the pre-tonic vowel was counted as lengthened in Table 2.6 was 0.8 of the length of the stressed vowel. As expected, pre-tonic lengthening was observed more frequently in Mizrahi speakers.

An additional measure of rhythm was added, based on the duration of vowels within the utterance. Following Ramus et al. (1999) and references there, I assume that perceived rhythm is not directly based on the duration of the interval between stressed syllables or the duration of syllables, as the original terminology used in the literature – “stress-timed” and “syllable-timed” languages – suggests. In fact, in survey after survey, inter-stress interval and syllable duration were found not to be correlated with perceived language rhythm (e.g. Dauer 1983). On the other hand, as Ramus et al. (1999) showed, the total duration of vocalic intervals in the sentence, divided by the total duration of the sentence, *does* correlate with rhythm perception. %V was therefore chosen as a second measure of rhythm. Since all speakers naturally produced the sentence as three distinct intonation phrases, each recording was divided into three: (I) halo? (II) jalom, (III) ani mitkafer-(et) legabej ha-dira be-rehov ha-ʕavoda. %V was calculated for each separately.<sup>14</sup>

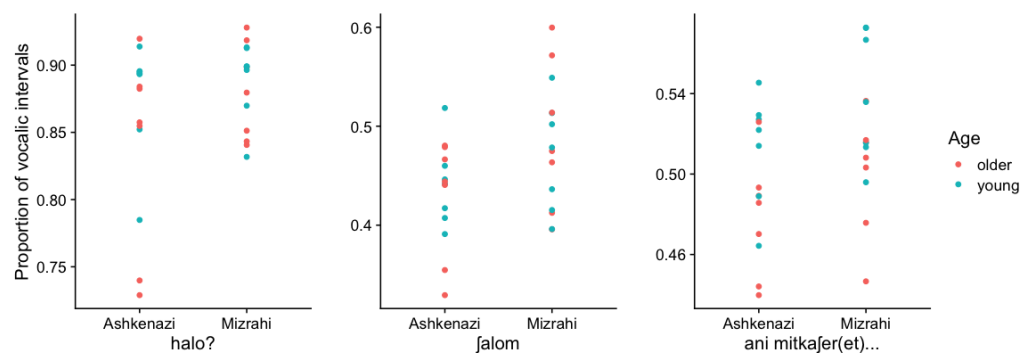


Figure 2.3: %V by intonation phrase (left to right: I-III), ethnic background and age.

The correlation between %V and the ethnic background of the speaker was significant in the

<sup>13</sup> Other words did not contain a suitable environment to measure pre-tonic lengthening, for one of the following reasons: (i) they were at a prosodic boundary, which might create a disproportionate final lengthening of the final syllable (“jalom”, “ʕavoda”, “halo”); (ii) they are function words (“ani”, “legabej”); (iii) The syllable structure of the word is not identical across speakers, due to an optional vowel deletion process ([be.re.ħov]~[ber.ħov]), or grammatical gender differences ([mitkafer]~[mitkafer+et]).

<sup>14</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 background; in addition, gender did not affect the results significantly in general and more particularly, in the relevant intonation phrase:  $t = 1.4791, Df = 25.937, p = 0.1512$ .

second and near-significant in the third intonation phrase, despite the relatively small sample of only 14 speakers per group (I:  $t = -1.49, Df = 19.46, p = 0.15$ ; II:  $t = -2.09, Df = 24.52, p = 0.047$ ; III:  $t = -1.8465, Df = 25.823, p = 0.07631$ ; see Figure 2.3).

In the first intonation phrase, there was a lot of individual-level variation. This is mainly due to two older Ashkenazi women who did not apply the common process of /h/-deletion (Matras and Schiff 2005, Gafter 2014). The relationship between %V and the ethnic background of the speaker remains significant when the entire utterance is considered ( $t = -2.1, Df = 25.9, p = 0.045$ ).

Pitch contours were classified as one of three contours speakers used: Falling, \*LH% (Rising), or \*HLH%. A contour was labeled as \*LH% if the boundary tone was the highest peak of the utterance. It was labeled \*HLH% if the highest nucleus peak was right after pitch reset. Overall, older speakers used a falling contour, while young speakers used either of the rising contours. More Mizrahi than Ashkenazi speakers used the \*HLH% contour, in both age groups. In addition, there was a significant correlation between the usage of the \*HLH% contour and %V: speakers who used the \*HLH% contour also tended to have a larger proportion of vocalic intervals within the speech stream (Pearson's correlation:  $t = 2.66, Df = 26, p = 0.013$ ).

**Procedure:** The experiment ran on *Minno*, a platform for running experiments online (Zlotnick et al. 2015). It proceeded as follows. Before the task, participants filled in a form collecting demographic data: age, gender, native language(s), main place of residence since childhood. In the main task, “Mizrahi” and “Ashkenazi” were presented at opposite ends of a 7-point scale, the mid-point being “I can’t tell”. Scores were set from -3 to 3. Participants were requested to wear headphones, listen to each speaker, and rate them according to how they sound. They could listen to each speech sample for as many times as they wanted, by pressing a button; there was no option to go back and re-rate a speaker once they had moved to the next screen. The speech samples were presented in random order. After the judgment task, subjects completed a short questionnaire about their demographics and attitudes.

#### **Questionnaire:**

A form at the end of the task requested for participants’ linguistic backgrounds and opinions on the following points (items appeared in random order):

1. I think that I sound: (7-point scale, identical to the one used in the classification task)
2. I produce [h] and [ʕ] in my speech. (Yes/ No/ Sometimes)
3. At least one of my parents produces [h] and [ʕ] in their speech. (Yes/ No/ I’m not sure)
4. At least one of my grandparents produces [h] and [ʕ] in their speech. (Yes/ No/ I’m not sure)

5. What's the best description of the environment you grew up in, in terms of the ethnic identities of people in your place of residence or at school? Mark according to whichever seems more relevant.
- Mizrahi majority
  - Ashkenazi majority
  - Very diverse
  - A majority of immigrants from the former USSR
  - A majority of immigrants from Ethiopia
  - Other
6. Which of the following definitions describes you best?
- I feel very attached to my Mizrahi identity
  - I feel relatively attached to my Mizrahi identity
  - I'm Mizrahi, but I don't really relate to these issues
  - I feel very attached to my Ashkenazi identity
  - I feel relatively attached to my Ashkenazi identity
  - I'm Ashkenazi, but I don't really relate to these issues
  - I'm of mixed origins and I'm connected with all the aspects of my identity
  - I'm of mixed origins and I'm more connected with my Mizrahi side
  - I'm of mixed origins and I'm more connected with my Ashkenazi side
  - My origins don't fall into these definitions and I feel they are absolutely irrelevant for me
  - My origins don't fall into these definitions and I feel culturally more related to Ashkenazis
  - My origins don't fall into these definitions and I feel culturally more related to Mizrahis
7. There are differences in pronunciation between Hebrew speakers from the center and speakers from the Periphery. (5-point scale between "absolutely agree" and "absolutely disagree", the midpoint being "I don't know")
8. Sometimes it is possible to know by the way a person speaks whether they come from the center or the periphery. (5-point scale between "absolutely agree" and "absolutely disagree", the midpoint being "I don't know")
9. The second generation of Mizrahi and Ashkenazi speakers sound different. (5-point scale between "absolutely agree" and "absolutely disagree", the midpoint being "I don't know")

10. The labels “Mizrahi” and “Ashkenazi” are no longer relevant for Israeli society. (5-point scale between “absolutely agree” and “absolutely disagree”, the midpoint being “I don’t know”)
11. Mizrahis and Ashkenazis have equal opportunities in society today. (5-point scale between “absolutely agree” and “absolutely disagree”, the midpoint being “I don’t know”)

Based on these questions, three scales had been created: Exposure to Variation, Attitude towards the Ethnic Cleavage and Acknowledgement of Variation. The scoring value for each response was based on a mid-point of indifference (= 0), e.g. “sometimes”, “I don’t know” or “not sure”, as in the classification task. That is, scores on either scale could be positive or negative. The score for “Mizrahi majority” in question 5 was 2 (maximum) and for “Ashkenazi majority” it was -2 (minimum). Diverse backgrounds, including a majority of immigrants from the USSR and Ethiopia, scored 1, and “other” was marked N/A. The scores for attitude questions were between -2 and 2, from “absolutely agree” to “absolutely disagree”.

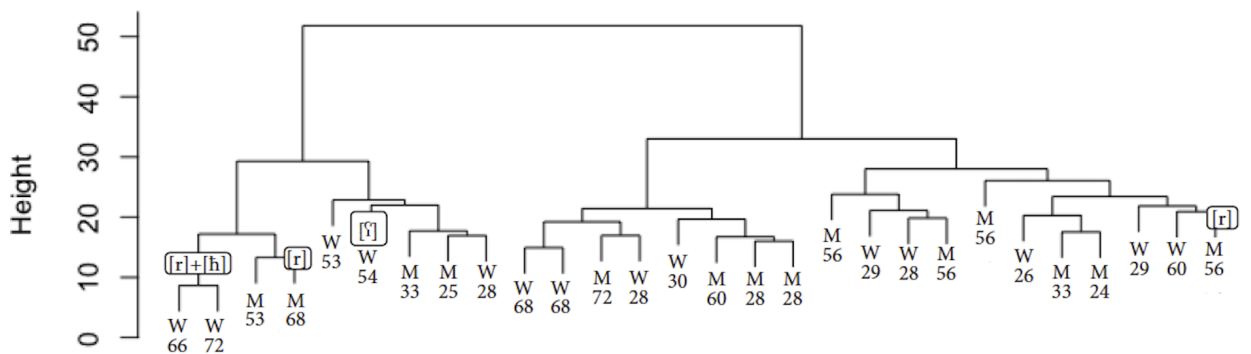
The answer to question 6 was split into two values: ethnicity, and *attitude towards ethnicity*. Responses to the latter got scores between 0-2, depending on the level of emotional involvement and regardless of ethnicity, e.g. “I feel very attached to my Mizrahi identity” and “I feel very attached to my Ashkenazi identity” both scored 2; “I don’t really relate to these issues” scored 0, regardless of the participant’s ethnicity.

The score for Exposure to Variation was calculated as the sum of responses to questions 2-5. These questions quite straightforwardly quantify the degree of self-reported exposure to variation. The score for Attitude towards the Ethnic Cleavage was calculated as the sum of responses to questions 9, 10, 11 and attitude towards ethnicity (extracted from 6 as described above). These questions target the participant’s belief regarding the contribution of *ethnicity* in particular to variation – be it linguistic or socio-economic. The score for Acknowledgement of Variation was calculated as the sum of responses to questions 7, 8 and the absolute value of 1. The logic of including 1 as an absolute value is that whether speakers perceive themselves to sound Mizrahi or Ashkenazi, they acknowledge that ethnically-conditioned variation exists.

## 2.2.2 Results and discussion

### 2.2.2.1 Inter-listener agreement

Hierarchical clustering was applied to the perceived ethnic identity ratings, using the `hclust` function in R (R Core Team 2017). This method initially assigns to each object (recording, in this case) its own cluster and then the algorithm proceeds iteratively, at each stage joining the two most similar clusters, until there is just a single cluster. 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



**Figure 2.4: Hierarchical clustering of perceived ethnicity, based on 105 ratings per recording. The branches containing known Mizrahi markers are labeled. Height represents euclidean distance.**

is on the left (the last, i.e., most recent, merge of the left subtree is at a lower value than the last merge of the right subtree). As visible in the dendrogram presented in Figure 2.4, the best choices for total number of clusters (that is, the least similar groups) are either 2 or 3, that can be labelled “Mizrahi”, “Ashkenazi”, “Ambivalent/Neutral”. The latter two are both less tight (as evident by their order from left to right) and closer to each other than the first.

As a means of validating the chosen clusters, Cronbach’s Alpha was computed for each. For the Mizrahi cluster (consisting of 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 ( $\alpha = 0.37, CI = 0.2 - 0.55$ ). Reliability is reduced with a two-cluster analysis, therefore the three-cluster cut is preferred. Table 2.7 includes the mean and standard deviation for each voice, divided into the clusters found by the algorithm.

Importantly, in the resulting Mizrahi cluster not all speakers used the marked Mizrahi consonants. In the Ashkenazi cluster, no speaker had used any known marker. Taking these results together, it can be stated that some speakers who do not use markers were perceived as sharing the ethnic identity of speakers who do use them, while others were not; or, in other words, there are significant differences in ratings of recordings that do not include known markers. This result suggests that there are other acoustic correlates on which listeners have relied in their ratings.

Usage of [h] reduced variation between listeners’ judgements (see standard deviation for 1 and 3). This finding is congruent with previous studies that found [h] to be consistently produced by Mizrahi speakers who maintain the distinction between the categories of [χ] and [h], contrary to the other markers, [ʔ] and [r]. In a recent production study, [ʔ] was shown to be used less consistently by the same speakers who maintain the [h]-[χ] distinction depending on style, and was additionally used by Ashkenazi speakers in some contexts (Gafter 2016a). Accordingly, the (only) speaker who produced [ʔ] (speaker 10) received mixed responses, as evident by the relatively high standard deviation. Similarly, [r] is used by speakers of various ethnic backgrounds, native

Cluster   label	Age	gender	Mean	SD	Markers
Mizrahi   1	72	W	2.63	0.82	[h] and [r]
Mizrahi   3	66	W	2.5	1.14	[h] and [r]
Mizrahi   4	68	M	2.04	1.4	[r]
Mizrahi   5	48	M	1.67	1.27	[r]
Mizrahi   20	25	M	1.58	1.32	
Mizrahi   26	28	W	1.23	1.34	
Mizrahi   28	33	M	1.21	1.36	
Mizrahi   10	54	W	1.05	1.8	[ʔ]
Mizrahi   15	53	W	0.48	1.54	
Ashkenazi   2	68	W	-2.24	0.95	
Ashkenazi   22	68	W	-1.98	1.37	
Ashkenazi   16	28	M	-1.69	1.33	
Ashkenazi   8	60	M	-1.65	1.35	
Ashkenazi   17	28	W	-1.66	1.4	
Ashkenazi   21	72	M	-1.52	1.47	
Ashkenazi   25	28	M	-1.33	1.31	
Ashkenazi   19	30	W	-0.84	1.39	
Neutral   7	60	W	-0.74	1.6	
Neutral   24	26	M	-0.65	1.5	
Neutral   12	24	W	-0.55	1.6	
Neutral   6	56	M	-0.50	1.9	[r]
Neutral   11	26	W	-0.44	1.6	
Neutral   9	56	M	-0.28	1.4	
Neutral   18	29	W	-0.22	1.6	
Neutral   27	33	M	-0.19	1.4	
Neutral   23	28	W	-0.12	1.5	
Neutral   13	24	M	0.21	1.3	
Neutral   14	56	M	0.17	1.8	

**Table 2.7: Mean rating and SD by cluster, age and gender. Existence of markers indicated in the rightmost column**

and non-native, including Arabic (associated with Mizrahi speakers) but also European languages (associated with Ashkenazi languages). There were three speakers who produced [r] and did not produce other segmental markers. Of them, two were rated as Mizrahi relatively reliably (speakers 4 and 5). The third (6), who is a descendent of Iraqi Arabic speakers, had the highest overall SD and was clustered by the algorithm with the “ambivalent/neutral” voices; many participants rated him as sounding Ashkenazi. There might have been other features of the recording that sounded more Ashkenazi, explaining why listeners tended to use both sides of the scale (contrary to stating they can’t tell, which was the case for many of the “ambivalent/neutral” voices).

Young speakers who showed no segmental variation were still in some cases classified Mizrahi or Ashkenazi. This was the case with 20, 26 and 28 – all from Mizrahi majority neighborhoods; and with 16, 17, 25 - all from Ashkenazi majority neighborhoods (see Table 2.5). Two of the voices classified as Ashkenazi were bilingual speakers of English, who spoke Hebrew at home and spent most of their childhood in Hebrew-speaking schools, but also spent several years at schools in the US. Their bilingualism is not likely to be the only reason for their rating as sounding Ashkenazi; a third young English-Hebrew bilingual, with a similar background but who grew up in a more diverse Hebrew-speaking environment (24), was not rated reliably (that is, he was put in the “ambivalent/neutral” cluster).

#### 2.2.2.2 *Acoustic correlates of the ratings*

Two linear logistic regression models were used in order to assess the contribution of various acoustic properties to the ethnicity ratings.<sup>15</sup> Each model was initially maximally specified, and fixed effects and interactions that did not significantly contribute to data likelihood (measured as the Bayesian information criteria) were gradually removed.

The first model included data from the older voices, since in this group there were speakers who used the known segmental markers. Based on previous studies, I assume that these markers govern ethnicity perception when they are available (see Section 2.1). The model initially included [r], [ʕ] and [h] as fixed factors (dummy-coded), in addition to Rhythm (%V, continuous) and Contour (\*LH%/ \*HLH%/ Fall, dummy-coded); Recording and Participant were included as random effects. Contour did not contribute to the model and was therefore removed. A variance inflation factor test was conducted, using the VIF function in R’s ‘Car’ package (Fox 2019). It raised a concern of multicollinearity for [r] (3.73) and [h] (3.57); Zuur et al. (2010) recommend a threshold of 3. Since there were more speakers who used [r] but not [h], and additionally, the two speakers who produced [h] had also used [r], [h] was removed, leaving the model summarized in Table 2.8.

This model confirms the observation from previous studies that the apical variant [r] and the pharyngeal [ʕ] contribute to ethnicity perception, and contributes the insight that rhythm may play a role as well. While [h] was removed from the model due to multicollinearity, previous studies

---

<sup>15</sup> Analyzing ordinal data using parametric tests has been under heated debate by statistics experts. I rely on the discussion in Sullivan and Artino Jr (2013) in analyzing the data as numeric.

	$\beta$	SE	df	t	p
Intercept	-13.68	2.4504	14	-5.58	0.0006
	2.14	0.34	14	6.22	0.0001
$\zeta$	1.91	0.61	14	3.11	0.008
Rhythm	23.95	4.6	14	5.2	0.0001

**Table 2.8: Model output for optimal model predicting ethnicity ratings for older speakers in Experiment 1. Significant factors are emphasized.**

have shown that it contributes to ethnicity perception; see Section 2.1. I do not intend to claim otherwise: this is merely a limitation of the current study, stemming from the particular features of the used recordings.

The second model included only young voices, none of whom used the known segmental markers. In this model, rhythm and contour were included as fixed effects, in addition to the random effects of Voice and Participant. Contour did not contribute to the model’s informativeness and was therefore removed. In this model as well, rhythm was significant ( $\beta = 22.75, SE = 6.37, df = 14, t = 3.57, p = 0.003$ ), such that a higher proportion of vocalic intervals correlated with more Mizrahi ratings.

The current experiment is a first quantitative exploration of the ethnic perception of phonetic variation in Modern Hebrew, beyond the known segmental markers. It set out to test not one specific variable, but a variety of features as they are found in natural 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, some of which might be correlated with rhythm. In other words, the experiment does not allow us to deduce a direct relationship between rhythm and ethnicity perception. Still, rhythm seems like promising direction for future studies of ethnicity perception in MH, for two reasons. First, rhythm was significantly correlated with perceived ethnicity for both age groups. Second, fine-grained distinctions in rhythm have been shown to correlate with vernacular perception in several other languages from different language families, as discussed in Section 2.1.2.

### 2.2.2.3 *The influence of listeners’ linguistic background and attitudes*

The main hypothesis regarding the potential influence of language ideology was that acknowledgement of the ethnic cleavage should be correlated with more acknowledgement of acoustic variation that stems from ethnicity. That is because, as discussed in Chapter 1, Hebrew speakers tend to conflate ethnicity and SES, such that variation is not necessarily perceived as related to the ethnic identity of the speaker. In order to measure participants’ willingness to apply an ethnic label to a given recording, a Response Spread measure was used, which calculated the standard deviation of responses to all voices by a given participant, i.e. how much a participant tended to be certain about their judgement of voices as Mizrahi or Ashkenazi (recall that each response was



coded as between -3 and 3). Low spread of ratings indicates that participants tended to rate all voices around zero (“I can’t tell”).

The Pearson correlation method was used with Response Spread, Exposure to Variation, Attitude towards the Ethnic Cleavage and Acknowledgement of Variation as factors. Of the ideological measures, the only significant correlation with Response Spread was Acknowledgement of Variation.

In addition, the two measures of Acknowledgement of Variation and Attitude towards the Ethnic Cleavage were correlated: participants who stated that ethnicity still played a role in the lives of Jewish Israelis also tended to explicitly acknowledge that there exists phonological variation (Table 2.9).

### 2.3 Conclusions

In Chapter 1, I showed that the literature and public discourse around variation in Modern Hebrew downplay the contribution of ethnicity to variation. The current chapter adds three main findings to the study of social perception of phonological variation in MH:

1. Socially meaningful phonetic variation in Modern Hebrew exists, above and beyond the known segmental markers.
2. There exists a distinct notion of an “Ashkenazi” speech style that is not (or at least, no longer) the unmarked neutral; some speakers seem to be perceived as “sounding Ashkenazi” in particular.<sup>16</sup>

<sup>16</sup>This point should be inspected more closely, since the experiment did not include a “neutral” option. I interpreted ratings around zero as stemming from unmarked speech; however, as with the ratings of the speaker 6 (discussed in p.53), it is in some cases the result of mixed cues, which send raters to opposite ends of the scale.

Parameter 1	Parameter 2	r	CI	95% CI	t	df	p
Spread	Acknowledgement of Variation	0.34	0.95	[0.16, 0.50]	3.67	103	<.002**
Spread	Attitude towards the Ethnic Cleavage	0.1	0.95	[-0.10, 0.28]	0.97	103	>.999
Spread	Exposure to Variation	0.18	0.95	[-0.02, 0.36]	1.81	101	.29
Acknowledgement of Variation	Attitude towards the Ethnic Cleavage	0.28	0.95	[0.10, 0.45]	2.99	103	.017*
Acknowledgement of Variation	Exposure to Variation	-0.005	0.95	[-0.20, 0.19]	-0.06	101	.999
Attitude towards the Ethnic Cleavage	Exposure to Variation	0.03	0.95	[-0.17, 0.22]	0.27	101	>.999

**Table 2.9: Correlation between the Spread of Ratings and views regarding variation.**

3. Speakers vary in their willingness to attribute phonological variation to ethnic identity and place of residence; a correlation was found between the level of explicit acknowledgement of variation and the levels of certainty in ratings (reflected in the spread of ratings along the scale).

Most previous studies on phonological variation in MH assume that the pharyngeals and /r/ were, and still are, the main dimension of ethnically-conditioned variation in Modern Hebrew. Variation outside of these variables is not often attributed to ethnicity. This view is implicit in studies that stress the difference between “Oriental” and “General” Hebrew (see discussion in Chapter 1), such that speakers who do not produce the marked sounds necessarily belong to the General Hebrew division, at least in terms of pronunciation. This view is sometimes stated more explicitly, as in Devens 1981, who explains the finding that some speakers in her sample were reliably rated as Ashkenazi: “Possibly this is based on an assumption that, if no overt signs of Middle Eastern background are present [meaning the pharyngeals and /r/, S.B], the speaker is most likely of European descent” (1981:27).

In the current experiment as well, the rating of some older speakers as Ashkenazi was almost as reliable as the rating of other older speakers as Mizrahi (see Table 2.7). This is to a certain extent compatible with the assumption that no markers → not Mizrahi → Ashkenazi, since listeners can more successfully rely on the markers with older voices. Additional support for the view that sounding Ashkenazi = sounding not Mizrahi, is that there is more variation in ethnicity perception (reflected in the standard deviation of ratings) when the known markers are absent. This was true regardless of whether voices were rated Mizrahi, Ashkenazi or ambivalent.

However, the current study also provides a first indication that speakers may, in fact, *sound Ashkenazi* – and not “general” or “neutral”. That is because even if we accept the interpretation of “no markers” = “not Mizrahi” for the older group of speakers, it cannot explain the differences within the young group, in which none of the speakers used the known markers. Further, it would be difficult to explain the fact that some young speakers were rated *Ashkenazi* rather than neutral.

What phonetic/phonological variables make listeners perceive young native Hebrew speakers as significantly different from each other in terms of ethnic identity? Rhythm was found to be significantly correlated with ethnicity perception in both age groups (Section 2.2.2.1). This preliminary result, albeit a correlation and not a causative link, suggests a previously unidentified prosodic marker of ethnic identity. It thus provides a new starting point for qualitative studies of variation in rhythm and pitch.

The third finding was that participants who explicitly acknowledged linguistic variation are those who expressed higher certainty regarding their ratings (i.e. had a wider “spread” of judgments), using the extreme edges of the scale and mostly avoiding the middle “I don’t know” option. This can be explained by the following hypotheses:

1. Listeners who are more experienced with linguistic variation are more familiar with its social meaning, and therefore more certain about their judgments.

2. There is an ideological divide between listeners who ascribe variation to ethnic identity and place of residence (who have a wide spread of judgments) and listeners who do not (narrow spread of judgments).
3. Speakers provided congruent judgements in both tasks. After completing the rating task, speakers reflected on how easy it was for them to perform it, and provided matching explicit judgements.

As evidence against the first hypothesis, it is worth pointing out that self-reported experience of participants with people of various ethnic identities did not correlate with the degree of certainty in ratings. That is, people who grew up in an ethnically diverse environment did not have a wider spread than people from ethnically homogeneous environments. Against the second hypothesis, note that the ideological measure of Attitude Towards the Ethnic Cleavage, which measured how strongly participants think that the ethnic cleavage still affects life in Israel today, correlated with Acknowledgement of Variation. Acknowledgement of Variation additionally correlated with Spread. Despite this, Attitude Towards the Ethnic Cleavage and Spread did not correlate, indicating that they are probably not associated. The third hypothesis therefore seems the most plausible explanation.<sup>17</sup>

Thus, the tentative answer to the second question presented at the opening of this chapter seems to be no – no direct ideological line can be drawn between the recognition of ethnicity as an organizing principle of life in Israel and the recognition of phonological variation as stemming from ethnicity.

---

<sup>17</sup> In order to assess it more directly, an alternative design would have some participants answer the questionnaire before the rating task. If the correlation between Spread and Acknowledgement of Variation is mediated between the order of the tasks, then it is likely to be correct.

## **Part II**

# **Exploring phonological representations of marked variants in Modern Hebrew**

### III

## Pharyngeal minds: different lexical representations for speakers of a single dialect

This chapter is based on a paper in progress, co-authored with Noa Bassel (The Hebrew University in Jerusalem) and Roey J. Gafter (Be'er Sheva University of the Negev).

### 3.1 Introduction

Linguistic variation is often expressed in the phonological inventory. In the case of Modern Hebrew, whereas some speakers maintain a phonemic distinction between pharyngeal and non-pharyngeal consonants, the majority of speakers have merged these segments and produce only the non-pharyngeal variants. This has long been observed for production (Davis 1984; Yaeger-Dror 1988; Lefkowitz 2004; Gafter 2016b, *inter alia*), and may lead to the intuitive conclusion that the majority of Hebrew speakers also lack a representation of pharyngeals in their phonological system, as they do not produce these sounds. However, recent work that combines insights from sociolinguistics and psycholinguistics demonstrates that dialect perception does not always align with dialect production. For example, Sumner and Samuel (2009) demonstrate that rhotic speakers from the New York City area that grew up exposed to the non-rhotic New York City variety are primed by non-rhotic stimuli in a fashion similar to non-rhotic speakers, as opposed to rhotic speakers without such exposure. Similarly, exposure to pharyngeals may affect the perception of Hebrew speakers to the extent that speakers who do not produce them maintain some mental distinction between these categories.

In this study, we focus on one of the two pharyngeal consonants in Modern Hebrew, the voiceless pharyngeal fricative [ħ]. For speakers who produce it, it contrasts with a dorsal fricative, [χ], whereas for most younger speakers, only the latter variant occurs.<sup>1</sup> Since the [ħ]-[χ] merger is an advanced change in progress in many Hebrew-speaking communities, it is quite common for parents who produce [ħ] to raise children who merge [ħ] and [χ] in production. Therefore, we distinguish not just between merged and non-merged speakers but between *three* groups of speakers:

---

<sup>1</sup> The production of the merged category might vary between a velar fricative, a uvular fricative, and a uvular trill (Gafter 2020).

Pharyngealizers, who produce both consonants in their speech; Pharyngeal Listeners, who were exposed to pharyngeals in their parents’ speech but do not produce them themselves; and Merged Listeners, who were exposed to and use only the merged variant [χ] (see Table 3.1). Pharyngealizers differ from the other two groups at the level of *production*, whereas Merged and Pharyngeal Listeners differ from each other at the level of the primary input they received during childhood, producing the same merged realization.

We conducted two experiments in order to investigate the status of [ħ] among these three groups: a categorization task, in which participants had to recognize [ħ] at the phonetic level; and a lexical decision task, which aimed to explore the encoding of [ħ] as independent or merged with [χ]. Pharyngealizers are, of course, expected to perform at ceiling in both tasks: they systematically distinguish between these categories in production, and must therefore represent them differently. As for the two Merged speakers groups, we hypothesize that Pharyngeal Listeners maintain an independent representation of /ħ/, while Merged Listeners maintain a single merged /χ/ category. Despite this difference at the lexical level, we assume based on previous sociolinguistic work that [ħ] is recognized by Merged Listeners at the phonetic level; that is because it is one of the two most salient markers of Mizrahi identity (see discussion below).

	Pharyngealizers	Pharyngeal listeners	Merged listeners
Exposure to pharyngeals	Yes	Yes	No
Production of pharyngeals	Yes	No	No

**Table 3.1: The three target groups of listeners in the study.**

### 3.1.1 Sociolinguistic background

As illustrated in Chapter 1, an important aspect of Jewish-Israeli society is the so-called ‘ethnic cleavage’ between Jews of European descent, known as *Ashkenazi*, and Jews of Middle Eastern and North African descent which have received the broad cover term *Mizrahi*. The persisting relevance of these social categories is reflected in the fact that Ashkenazis on average have higher earnings and education compared to Mizrahis (Lewin-Epstein and Cohen 2019; Dahan et al. 2003), and are considerably over-represented in the local cultural and economic elite (Sasson-Levy 2013).

This ethnic distinction also has consequences for linguistic variation. Modern Hebrew emerged as a spoken language in the late 19th and early 20th century, and evolved as an amalgamation of various earlier traditions (as Hebrew was retained as a liturgical language by many Jewish communities; Morag 1990). A central phonological difference between the Ashkenazi tradition, based in Central and Eastern Europe, and traditions which developed in Muslim countries, lies in the status of two historical pharyngeal segments: a voiceless pharyngeal fricative ([ħ]) and a voiced pharyngeal approximant ([ʕ]). The traditions developed in Muslim countries – which later became known as Mizrahi – retained the pharyngeal consonants, in accordance with earlier strata

of the Hebrew language, whereas Ashkenazi traditions merged them with non-pharyngeals ([χ] for [ħ], and /ʔ/~/Ø/ for [ʕ]). The divergent traditions translated into a sociolinguistic marker in Modern Hebrew, with the retention of the pharyngeals (as a set of the two) becoming the linguistic feature most widely associated with Mizrahis (Matras and Schiff 2005; Gafter 2016a). However, scholars of Hebrew sociolinguistics are generally in agreement that the production of pharyngeal consonants have been undergoing a steady decline in most Mizrahi communities (Blanc 1968; Yaeger-Dror 1988; Davis 1984; Lefkowitz 2004; Gafter 2016b). The relevant aspect for our study is that second generation native Hebrew speakers do not distinguish [ħ] and [χ] in production, and produce some variant of [χ] for both categories. This change in progress is likely an outcome of the fact that retention of the pharyngeals became a stigmatized feature, following from the socio-economic inequality between Ashkenazis and Mizrahis (Ben Tolila 1984; Gafter 2016a).<sup>2</sup>

Following Blanc (1968), sociolinguistic descriptions of Hebrew have traditionally spoken of two main sociolects: ‘Mizrahi Hebrew’ as opposed to a so-called ‘General Israeli’, representing the alleged unmarked speech of Ashkenazis (see for example Matras and Schiff 2005; Yaeger-Dror 1988; Davis 1984). The main phonological differences between these two varieties are the aforementioned pharyngeals, as well as divergent realizations of the Hebrew rhotic. However, such a binary view does not adequately describe the extent of linguistic variation, nor the enormous heterogeneity among both Mizrahis and Ashkenazis. Rather, a more apt description may be a continuum ranging between typically Mizrahi and typically Ashkenazi variants, with most native speakers falling along some point in between (Lefkowitz 2004).<sup>3</sup> Indeed, speakers who produce pharyngeals often do so variably, and a speaker may exhibit a mixture of these features, e.g. produce the distinction between [ħ] and [χ] but not [ʕ], etc. Further complicating the picture is the fact that the label ‘Mizrahi’ covers many different historical communities with distinct cultures and immigration histories, as well as linguistic patterns. Among Mizrahis, Jews of Yemenite descent are typically seen as those most likely to retain the pharyngeals (Davis 1984). This is not merely a linguistic stereotype: Gafter shows that the decline of pharyngeals has progressed more slowly among the Yemenite community in the town of Rosh Ha’ayin, comparing to a mixed sample of Mizrahis in neighboring Tel Aviv (Gafter 2016b). Accordingly, the groups of Pharyngeal Speakers and Listeners who participated in the current study are of Yemenite descent.

/ħ/ and /χ/ are associated with two distinct letters, ה and ח respectively. Thus, whereas speakers’ access to phonetic evidence for the distinction varies greatly based on the dialect they were exposed to, all literate Hebrew speakers are taught to differentiate the two letters in spelling (though teachers generally do not expect this to be reflected in students’ pronunciation, as they

<sup>2</sup> One could argue that phonological markedness may play a role here as well, as [ħ] is an uncommon sound in the world’s languages. However, no known dialect of Arabic, which similarly includes both of these categories, includes a merger between them. The merger can straightforwardly be traced back to the liturgical tradition of Ashkenazi Jews, who produce [χ] for both categories when reading in Hebrew.

<sup>3</sup> This is a gross simplification, as new waves of immigration, most notably from the former USSR and Ethiopia, have brought additional social and linguistic influences that cannot be located neatly between these opposing poles. However, the Mizrahi-Ashkenazi opposition is still very salient in Israeli discourse, and as such it is a valuable sociolinguistic resource with respect to which speakers can position themselves.

Hebrew Letter	Non-Merged	Merged
ה	[h]	[χ]
כ	[χ]	

**Table 3.2: Merged and non-merged realizations.**

often merge them themselves). This orthographic distinction may facilitate the acquisition of a perceived distinction between these categories, even for Merged Listeners (see Chéreau et al. 2007; Grainger and Ziegler 2011; Ziegler et al. 2004). In theory, Merged Listeners could rely on orthography to develop separate representations for ה and כ, despite their identity at the phonetic level. At the same time, the relationship between orthography and production is somewhat more complex than is suggested in Table 2, because the letter כ is in fact ambiguous between [χ] and a velar stop [k]. There is therefore no strict one-to-one correspondence between orthography and phonology in the case of [χ] in either variety.

### 3.1.2 Perception of merged categories

In order to understand what happens when merged speakers hear [h] – a consonant that exists in their native language but not in their own native variety of said language – we can consider the difficulty caused to L2 speakers by unfamiliar sounds in a new language. It has long been known from studies on L2 acquisition that the difficulty to produce L2 sounds that do not exist in one’s L1 is rooted not only in articulatory factors, but also in perception. Cases of perceptually-merged L2 sounds occur even with highly proficient speakers. For example, Parllier et al. (2001) found that Spanish-dominant early bilinguals experienced repetition priming for Catalan /e/~ε/ minimal pairs, contrary to native Catalan speakers. This was interpreted as evidence that L1 Spanish speakers not only produce these categories as merged, but also perceive such pairs as homophones. Using the same paradigm, Dufour et al. (2007) found that southern French speakers, unlike speakers from the Paris area, perceive /ε/~e/ minimal pairs to be homophones, showing that dialect-specific mergers can also induce significantly different lexical representations.

These studies contribute the insight that highly proficient speakers – in the case of French, native speakers – can still vary significantly from each other in terms of their lexical encoding. The nature of this difference requires more attention; there is a growing body of evidence suggesting that despite these seemingly symmetrical results (/ε/ words prime their minimal pair /e/ words and vice versa), lexical encoding is influenced by the phonetic resemblance of each category to a native L1 category of the listener. For example, Weber and Cutler (2004) found that L1 Dutch speakers who are highly proficient in English experience competition from /ε/ items upon hearing /æ/. That is, words that contain /æ/, such as *panda*, briefly activate words that contain /ε/, e.g. *pencil*. On the other hand, no competition effect was found in the opposite direction – upon hearing *pencil*, *panda* was not activated. This suggests that the category that is phonetically closer to an L1 category – in the Dutch-English case, /ε/ – is encoded more accurately in the lexicon, creating a more selective



pattern of activation (in this case, reflected in the eye movements of participants in a Visual World Task). In other words, the category that is phonetically similar to the native category (hereafter the “familiar” category) is more lexically constraining than the “new” category, which is perceived and lexically encoded as a “bad exemplar” of the old category.

Two additional studies support this view. L1 English speakers have a difficulty producing the distinction between front and back round vowels that are found in German (“familiar” category = back round vowels), and between singleton and geminate consonants in Japanese (“familiar” category = singleton consonants). Darcy et al. (2013) found that while intermediate and advanced learners are almost on a par with native speakers in terms of their phonetic decoding (tested in an ABX task), their performance in a lexical decision task in which these sounds were switched is significantly worse. Again, performance with switched “familiar category” items was better: real words were more likely to be accurately accepted if they included the familiar category, and non-words were more likely to be rejected if they included the new category. For example, the word *honig* ‘honey’ includes a back round vowel, while *könig* ‘king’ includes a front round vowel. L1 English learners of German more often correctly rejected a switch from the familiar to the new category (e.g. *\*hönig*) compared with the opposite direction of the switch (*\*konig*); they were also more likely to identify *honig* as a word, compared with *könig*. Melnik and Peperkamp (2019) showed similar findings with L1 French speakers learning English – performance on the /h/~/Ø/ distinction was better with [Ø]-words (e.g. *officer*) compared with h-words (*husband*), and category-switched non-words with the added new category (*\*hofficer*) were easier to reject than category-switched non-words with the familiar category (*\*usband*).

While the case of mergers between dialect-specific L1 categories, such as the Hebrew [h]-[χ] merger, has clear similarities to L2 learning, there are also important differences: unlike L2 learners, sounds of another dialect may be new, but the lexical items, of course, are not, as different dialects of the same language share most of the words. Therefore, speakers of a merged dialect receive direct evidence from their native lexicon during acquisition that there is *no difference* between the categories they might be facing when moving to a non-merged environment. For example, in the *pin-pen* merger in the US (Labov et al. 2008), a learner who had been exposed only or mostly to merging speakers during childhood is likely to perceive *pin* and *pen* as homophones (Conrey et al. 2005).

At the same time, the ‘new’ phonetic category (in the sense of Darcy et al. 2013) faced by merged native speakers is not necessarily entirely new: in many cases, speakers have some early exposure to the phonetic categories of non-merged varieties of their own language, which is often not the case with L2 learners. Often, phonological variation of this sort bears social meaning. That is certainly the case with the Hebrew pharyngeal fricative [ħ], which is highly stereotyped (Gaftor 2019). In this case, the ‘new’ phonetic category may actually be quite salient as a feature of a social group, but not necessarily as a lexical feature.

### 3.2 Experiment 2: Discrimination between [ħ] and [χ] in speakers with merged productions

There is ample evidence that [ħ] is a stereotyped sound in Modern Hebrew (Lambert et al. 1965; Davis 1984; Ben Tolila 1984; Gafter 2016a, among many others). While it is quite clear that usage of [ħ] is registered by merged listeners (Gafter 2016b), it has not previously been tested whether they can reliably distinguish these sounds from each other within the speech of a single speaker. That is, it is possible that Merged speakers register that the sound produced by Pharyngealizers for their merged [ħ]~[χ] category is generally different from the sound that they produce themselves and are used to hearing, but that they do not note the distinction that pharyngealizers make between [ħ] and [χ].

#### 3.2.1 Methods

In order to test whether Merged speakers can perceive the difference between [ħ] and [χ], we conducted a short online discrimination task using the PsychoPy platform (Peirce and MacAskill 2018). We used non-words produced by a native Hebrew speaker, half of which included [ħ] and half included [χ]. The recorded speaker is a Pharyngeal listener, i.e. his parents produced the distinction as well as some other characteristics of the dialect, but he does not in his daily life. However, he was comfortable with producing pharyngeals and a trilled [r], the stereotyped phonological features of the Mizrahi variety. Participants were instructed to press the letter key ם for non-words that include [χ], and the letter ן for non-words that include [ħ]. That is, the choice of keys utilized the orthographic knowledge of Hebrew speakers that [χ] is represented by the first letter, and the pharyngeal fricative is represented by the second.

##### 3.2.1.1 Participants

Seventy Hebrew speakers volunteered to participate in the experiment. Eighteen were Pharyngealizing speakers, all native Palestinian Arabic speakers with high proficiency in Hebrew. These speakers are expected to be on a par with native Pharyngealizing Hebrew speakers since they share the phonemic distinction.<sup>4</sup> Ten participants were native Hebrew speakers who had at least one pharyngealizing parent (“Pharyngeal listeners”). Forty-two were native Hebrew speakers who are second generation Merged speakers. None of the participants had reported hearing problems, dyslexia or ADHD.

##### 3.2.1.2 Stimuli

Thirty-eight pseudo-nouns, complying with Hebrew morphology and phonotactics, were used in the task, in addition to 6 practice trials. Half of the items included [ħ] and half included [χ].

---

<sup>4</sup>Native Palestinian Arabic speakers were more accessible to our online recruitment than older Mizrahi speakers, hence they were chosen as a control group. The categories are distinct in every known Arabic dialect.

### 3.2.1.3 Data analysis

For the purpose of the current analysis, Pharyngealizing speakers and Pharyngeal Listeners were considered as one group, that is expected to perform at ceiling. The fixed effects were therefore Participant Group (Pharyngealizing parent/No pharyngealizing parent) and Condition ([h]/[χ]). Accuracy and RT models included item and participant as random intercepts. Only correct responses were used in the RT model.

### 3.2.2 Results

The unmarked category [χ] was responded to less accurately ( $\beta = -1.13$ ,  $SE = 0.27$ ,  $z = -4.2$ ,  $p < 0.0001$ ) and identified slower than its marked counterpart [h] ( $\beta = 0.08$ ,  $SE = 0.03$ ,  $Df = 38.53$ ,  $t = 2.34$ ,  $p = 0.02$ ).

The Participant Group also significantly affected categorization of [h]~[χ], such that listeners with Pharyngealizing parents were more accurate and marginally faster (Accuracy:  $\beta = 1.33$ ,  $SE = 0.64$ ,  $z = 2.08$ ,  $p = 0.04$ ; RTs:  $\beta = -0.11$ ,  $SE = 0.05$ ,  $Df = 76.96$ ,  $t = -2.07$ ,  $p = 0.04$ ). No significant interactions were found. Figure 3.1 is a visual summary of the results.

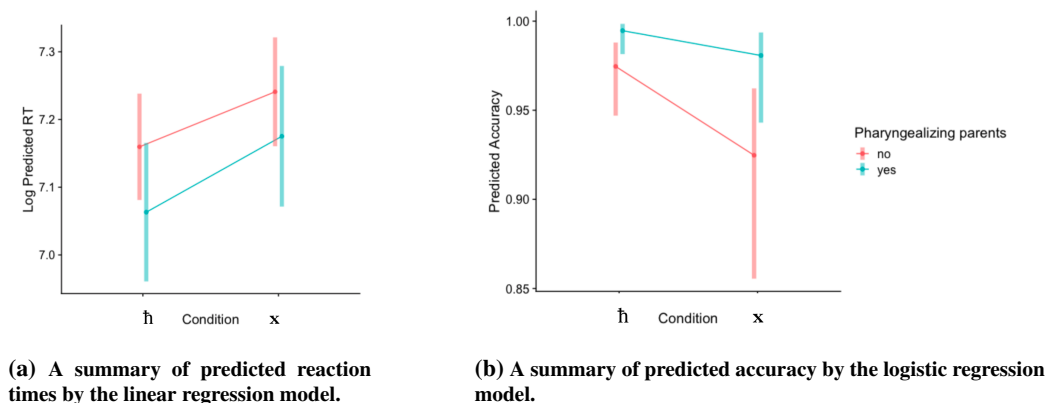


Figure 3.1: Results of Experiment 2.

### 3.2.3 Discussion

The results of Experiment 2 demonstrate that Merged speakers can discriminate between [h] and [χ] at a near-native level: the accuracy rates of Merged listeners were lower compared with speakers from a Pharyngealizing background, but still very high. Both groups had more errors and longer RTs in [χ] trials, indicating that it is easier to categorize the socially marked sound [h].

We can now clear off the table the option that Merged listeners cannot reliably perceive the difference between [h] and [χ] at the phonetic level. This is in line with previous indications that

[h] is salient: it is stereotyped (socially marked sounds are expected to be salient for listeners); Merged speakers are overtly taught in school that there are two underlying categories involved in the production of Pharyngealizers (direct instructions enhance attention); and this distinction is also supported by orthography and by frozen phonological alternations that occur around pharyngeals as well as the alternation between [k] and [χ] for the uvular category (orthography and phonology support the distinction).

### 3.3 Experiment 3: A switched-category lexical decision task

Experiment 3 aimed to explore the lexical representations of Merged listeners, in addition to the two other groups: Pharyngealizers and Pharyngeal listeners, i.e., Merged speakers who have been exposed to pharyngeals since childhood. We used a lexical decision task in which the target categories had been switched with each other in some of the items, predicting different outcomes in terms of grammaticality based on the linguistic background of the listener. We used stimuli recorded by two speakers, introduced in separate blocks: one speaker merged the categories in his speech (a Merged variety voice) and the other did not (a Pharyngealizer voice).

We hypothesize that Merged listeners only have one category in their lexicon – /χ/, contrary to the two other groups. Therefore, we expect that regardless of whether items include an etymologically correct h or the switched-category  $\chi \rightarrow h$ , reactions should be similar. That is because the mechanism that maps surface [h] into the lexical category of /χ/ is the same in both cases, as there is no independent representation for [h] in the lexicon. As for the  $h \rightarrow \chi$  condition, the result of this switch is the same merger occurring in their own production, and is thus predicted to be processed easily.

The status of switched-category productions is more complex for listeners from pharyngealizing backgrounds. The  $\chi \rightarrow h$  direction of the switch does not occur naturally in either variety, and is therefore expected to be rejected by speakers from a pharyngealizing background. Items produced in the natural direction of the merger, i.e. the switched-category  $h \rightarrow \chi$  condition, are predicted to be lexically valid for this group as well, since this is the invariably produced category in the Merged variety (which is also the more common variety). However, merged productions are unexpected from a Pharyngealizing speaker, and this identity inconsistency may affect the processing of such items.

The perceived identity of a speaker was shown to affect various aspects of sense-making, including ambiguity resolution (Walker and Hay 2011; Cai et al. 2017), acceptance of divergence in production (Weatherhead and White 2018) and compensating for known differences in phoneme boundaries (D’Onofrio 2018b; Lev-Ari et al. 2019). For example, Walker and Hay (2011) found that words that are more likely to be used by older speakers according to a corpus were more quickly recognized when uttered by an old voice compared with a young voice, and vice versa. Cai et al. (2017) demonstrate that listeners tend to understand *bonnet* as a type of hat when uttered with an American accent, and as a car hood when the accent was British. That is, listeners have

different expectations regarding the words that a given speaker might use, and even regarding the particular meaning that the speaker intended in the case of ambiguous words. Given this, we might expect that bi-dialectal speakers will be affected by the persona of the speaker, such that merged production items should raise more difficulty when produced by the Pharyngealizing voice compared with the Merged voice.

Table 3.3 summarizes our predictions for the pattern of responses to switched-category items by each participant group. The mark “V” indicates no expected difficulty in the processing of items under the given condition. The mark “X” indicates expected rejection due to ungrammaticality – this is predicted only for items that do not occur in either variety, i.e.  $\chi \rightarrow \text{h}$  items. The mark “?” was used for items that are lexically valid, but inconsistent in terms of the identity of the speaker. While similar manipulations have been shown to affect various behavioral measures, as described above, it is unclear whether acceptance rates may be affected as well.

Group	$\text{h} \rightarrow \chi$ (Dialect-Inconsistent)	$\chi \rightarrow \text{h}$ (Pathological)
Merged	V	V
Pharyngeal listener	?	X
Pharyngealizer	?	X

**Table 3.3: Prediction for pattern of acceptance of switched-category stimuli produced in the Pharyngealizing speaker’s voice, for the three groups of participants in Experiment 3.**

### 3.3.1 Methods

#### 3.3.1.1 Participants

Seventy-eight native speakers of Modern Hebrew volunteered to participate in the experiment, divided between Pharyngealizing speakers (24), Pharyngeal listeners (25), and Merged listeners (29). The criteria for inclusion in the second group was that the participant merged the two categories in speech, and that a primary caregiver of the participant (a parent) was a Pharyngealizing speaker. The criterion for inclusion in the merger group was that the participant did not have early exposure to pharyngeal dialects (including extended family and friends). Note that speakers in the two latter groups produce the same merged variant. None of the participants in this experiment participated Experiment 2.

Recruitment was done through the authors’ friends and family. Participants were not asked about their linguistic background before the task; if after the exit questionnaire it turned out that the participant did not fit neatly into one of the groups, their data was discarded (e.g. a merged listener was discarded because they had a Pharyngealizing grandparent; a suspected Pharyngeal listener didn’t in fact have a Pharyngealizing parent). The mean age of participants was 58 in the Pharyngealizing speakers group (SD = 10), 31 in the Pharyngeal listeners group (SD = 10) and 34

in the Merged listeners group (SD = 16).<sup>5</sup>

Participants had no reported hearing problems. Six participants from the Pharyngeal listener group and one from the Pharyngealizing group reported being diagnosed with ADHD. Since the task requires some focus, this is not ideal; however, since these are the groups of participants who are expected to perform at ceiling, any effect it may have on the result is in the opposite direction of our hypothesis. Eight Merged listeners and five Pharyngeal listeners reported having a second native language (Russian or English). Three Pharyngealizing speakers were native speakers of Yemeni Arabic, in addition to Modern Hebrew.

### 3.3.1.2 *Stimuli*

Fifty-six Hebrew words served as critical items, of which 28 included [χ] and 28 [ħ] (none of the items contained both categories). Critical items were carefully selected such that the status of [ħ]/[χ] could not be determined on the basis of phonological alternations. For example, word initial [ħ] was not used, since the etymological [χ] cannot appear in this position phonotactically: word-initial [χ] surfaces as [k]. Additionally, 56 nouns which do not include either of the categories were used as distractors (Baseline condition in Table 3.4), as well as 112 pseudo-nouns which were compatible with Hebrew nouns phonotactically and prosodically, and included [χ] and [ħ] to the same extent as the real words in the experiment: 28 with [χ], 28 with [ħ] (the same non-words used in Experiment 2) and 56 with neither [χ] nor [ħ].

All target words were nouns. We controlled for the position of the target sound (coda/intervocalic), Concreteness (physical item/ not a physical item), number of syllables (2 or 3) and frequency; these features were evenly distributed between four mini-blocks. Each mini-block contained 7 [χ] items, 7 [ħ] items, 14 distractors and 28 non-words.

The stimuli were recorded twice, by two cisgender men who are native speakers of Hebrew, aged 30 and 34. The first is a native Merged speaker, and his recordings included the unmarked variants of the pharyngeals. The second was the same speaker who recorded the stimuli for Experiment 2. In one version of his recording of the stimuli, the pharyngeals were used consistently and correctly. In addition, he recorded a second version of each [ħ]/[χ] item: a “pathological”/“inconsistent” one, respectively. We use the label “pathological” to denote the cases in which original [χ] was replaced with [ħ], since this is a variant of the lexical item that does not exist in either dialect. We use the label “dialect-inconsistent” to denote the cases in which original [ħ] was replaced with [χ], since this is on the one hand the unmarked, merged category produced by Merged speakers; and on the other, it is not consistent with the dialect of the pharyngealizing speaker.

---

<sup>5</sup> Since Pharyngealizing speakers in society are generally older (due to the decline in usage of pharyngeals in younger generations) compared with the mean age of Pharyngeal Listeners (being the 2nd generation), we intentionally set out to test Merged listeners of all ages, in order to have a wide variety of age representatives in this group that would match Pharyngealizing speakers on the one hand and Pharyngeal Listeners on the other. Hence the larger standard deviation in mean age in this group. Due to a technical issue, age data is missing for 5 participants (3 Merged listeners and 2 Pharyngealizing speakers).

The experiment proceeded in three blocks, made up of 1-2 mini-blocks: one block in the Merged voice (one mini-block), another in the Pharyngealizing voice (one mini-block), and a third, again in the Pharyngealizing voice, included “pathological” and “inconsistent” trials (two mixed mini-blocks, of which one was reliable). The order of the first two blocks was randomized, but the third – i.e., the block that included Pathological and Inconsistent items – was always presented last. This was done in order to make sure that the Pharyngealizing speaker could initially be characterized as reliable (i.e. not as someone who makes production errors, from the point of view of non-merged speakers). The critical items were rotated between the conditions in a Latin square design, using the four mini-blocks such that participants listened to an item exactly once – produced by the Merged speaker, the consistent Pharyngealizer or the Pathological/Inconsistent Pharyngealizer. Sample items of each condition are provided in Table 3.4.

Voice	Baseline	/χ/	/ħ/	Inconsistent [χ] (ħ→χ)	Pathological [ħ] (χ→ħ)
Merged	atsits	mexonit	βexov	N/A	N/A
Pharyngealizer	ʃatsits	mexonit	rehov	?rexov	*mehonit
	‘pot’	‘car’	‘street’		

**Table 3.4: Sample stimuli in the Merged and Pharyngealizing voices. Note that in addition to variation in the production of target sounds, the speakers differ from each other in the production of /r/ and /ʃ/, in accordance with their spoken variety.**

### 3.3.1.3 Procedure

The experiment was run on PsychoPy (Peirce and MacAskill 2018). Due to the outbreak of Covid-19, we had to replace offline testing with an online version after roughly 2/3 of the data were collected. We asked participants to wear headphones throughout the experiment; in the offline version, we used Sony MDRZX100 ZX Series Stereo Headphones, on a 13-inch MacBook Air computer with 2.2 GHz Intel Core i7. Due to the less-controlled environment in which the online experiment was conducted, we added Running mode (Online/Offline) as a factor to both RT and Acceptance rates models. Additionally, as a sanity check, we analyzed the offline data again separately to make sure that the same trends are found in the data. No differences in trends were observed.

### 3.3.1.4 Data analysis

Only real words were analyzed. “Real words” include words produced in the Pathological condition (χ→ħ) and in the Inconsistent condition (ħ→χ in the Pharyngealizing voice). One item from the Baseline condition ([ʃiton] ‘newspaper’) was removed due to low acceptance rates (68%). All other real words were accepted in over 70% of all trials, excluding the Pathological condition. As acceptance rates for Pathological items rely on perceived well-formedness, which was predicted

to be different across groups, no Pathological item was removed from the analysis, despite their relatively low overall acceptance rates.

The accuracy criterion for participant exclusion was below 70% success with real word stimuli produced in the Merged voice. This condition was chosen as the performance test since the majority of daily input from the media is produced in this variety, and not accepting words produced this way probably stems from a general issue with the task. Three participants were excluded based on this criterion (2 Pharyngealizing speakers and 1 Pharyngeal listener).

Analyses were conducted using the lmerTest package (Kuznetsova et al. 2017) in the R software environment (R Core Team 2017). A logistic model was used for accuracy/acceptance data, and a linear model for RT data. The models included random intercepts for participants and items, and the following fixed factors and their interactions:

- Voice (Merged/ Pharyngealizer) was sum-coded;
- Condition was dummy-coded, the baseline being items that included No [h]/[χ]; the conditions compared with the baseline were [h] items, [χ] items, Inconsistent and Pathological, the latter two being relevant only for the Pharyngealizer voice blocks;
- Group was contrast-coded, with two planned comparisons: Merged listeners vs. Pharyngealizers and Pharyngeal listeners, and Pharyngeal listeners vs. Pharyngealizers.

Two additional fixed factors for which we did not consider interactions were Running mode (Offline/Online) and age. The bobyqa optimizer was used to allow the models to converge (Bates et al. 2007).

### 3.3.2 Results

#### 3.3.2.1 Acceptance rates

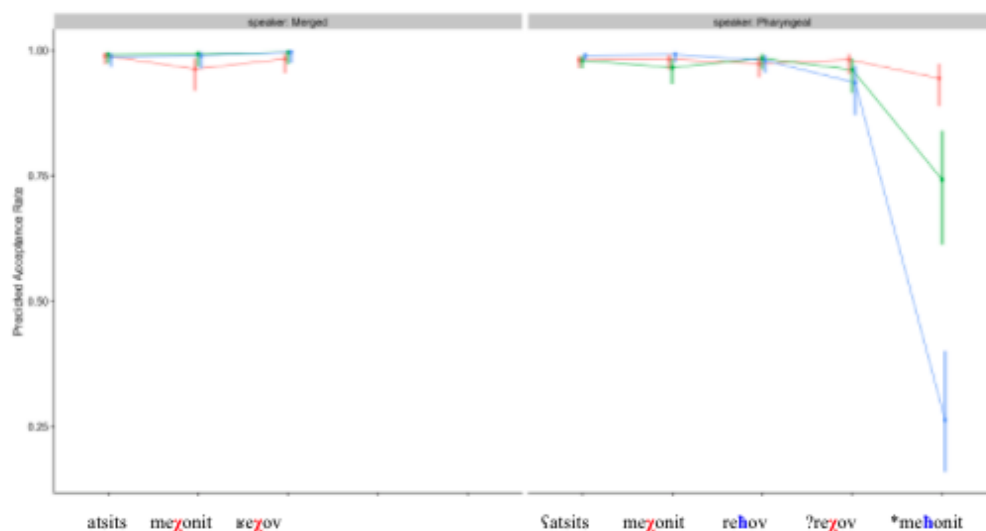
The acceptance rates model did not converge when Age was included as a fixed factor, therefore it was removed, leaving the final model as follows ( $N = 8325$ ):

Acceptance rates model = response  $\sim$  Condition \* Voice \* Group + running mode + (1 | participant) + (1 | item)

The logistic regression analysis revealed a main effect of Condition, with the Pathological and Inconsistent items conditions yielding significantly lower acceptance rates (Pathological:  $\beta = -3.17$ , SE = 0.28,  $z = -11.44$ ,  $p < 0.0001$ ; Inconsistent:  $\beta = -0.84$ , SE = 0.32,  $z = -2.66$ ,  $p < 0.008$ ).



Interactions were found between Participant Group and Condition, such that the Pathological condition affected listeners from pharyngealizing backgrounds (Pharyngealizers and Pharyngeal listeners) more than it did the Merged listeners group, and the Pharyngealizers more than Pharyngeal listeners (Merged vs. Pharyngealizers + Pharyngeal listeners:  $\beta = -1.04$ , SE = 0.13,  $z = -8.13$ ,  $p < 0.0001$ ; Pharyngeal listeners vs. Pharyngealizers:  $\beta = -1.38$ , SE = 0.21,  $z = -6.51$ ,  $p < 0.0001$ ). The same was true with the Inconsistent condition (Merged vs. Pharyngealizers + Pharyngeal listeners:  $\beta = -0.44$ , SE = 0.17,  $z = -2.6$ ,  $p < 0.009$ ; Pharyngeal listeners vs. Pharyngealizers:  $\beta = -0.61$ , SE = 0.27,  $z = -2.29$ ,  $p < 0.02$ ). Finally, there was a triple interaction between Voice, Condition and Group, such that Merged listeners made more errors with [χ] items in the Merged voice ( $\beta = 0.29$ , SE = 0.14,  $z = 2.03$ ,  $p < 0.04$ ). Figure 3.2 is a visual summary of the model, using the emmeans package of R (Lenth et al. 2018).



**Figure 3.2:** A summary of the logistic regression model of acceptance rates in Experiment 3. Blue: Pharyngealizers, Green: Pharyngeal listeners, Red: Merged listeners. The Merged voice condition is on the left panel; note that the inconsistent and pathological conditions are irrelevant for this voice, since his productions are merged. On the right – the Pharyngealizing voice. Conditions from left to right within each panel: Baseline, /χ/, /h/, Dialect-Inconsistent [χ], Pathological [h].

### 3.3.2.2 Reaction times

Only correct responses were considered. RTs were measured from the onset of the stimulus and log-transformed before being entered into the model. Responses below 0 or above 5 seconds were removed (0.5% of the responses), and responses of above 2.5sd of the participant’s mean were replaced by the mean+2.5sd (7 trials). The final analysis included 7877 trials, and the model was:

RT model = logRT  $\sim$  Condition \* Voice \* Group + age + running mode + (1 | participant) + (1 | item)

The model revealed a main effect of Condition, such that / $\chi$ /, /h/ and Pathological items were responded to more slowly compared with the Baseline (/ $\chi$ /:  $\beta = 0.03$ , SE = 0.01,  $df = 176.6$ ,  $t = 2.06$ ,  $p < 0.05$ ; /h/:  $\beta = 0.057$ , SE = 0.015,  $df = 176.2$ ,  $t = 3.71$ ,  $p < 0.001$ ; Pathological:  $\beta = 0.14$ , SE = 0.022,  $df = 281.9$ ,  $t = 6.215$ ,  $p < 0.0001$ ). The main effect for / $\chi$ / and /h/ is likely to be due to the overall high rate of items (words and non-words alike) that include [ $\chi$ ] or [h] – two thirds of the stimuli. Particularly due to the non-words, might have created a general carefulness around trials containing these sounds. The lack of a similar effect for the Inconsistent condition, which also contains [ $\chi$ ], is due to interactions in opposite directions with Group, as discussed below.

Another main effect was that of Voice: the Merged voice was responded to faster than the Pharyngealizing voice ( $\beta = -0.025$ , SE = 0.011,  $df = 147.7$ ,  $t = -2.281$ ,  $p < 0.024$ ).

Mimicking the acceptance rates, interactions were found between Group and Condition, such that the Pathological condition affected listeners from pharyngealizing backgrounds more than it did the Merged listeners group, and the Pharyngealizers more than Pharyngeal listeners (Merged vs. Pharyngealizers + Pharyngeal listeners:  $\beta = 0.033$ , SE = 0.008,  $df = 7702$ ,  $t = 4.057$ ,  $p < 0.0001$ ; Pharyngeal listeners vs. Pharyngealizers:  $\beta = 0.046$ , SE = 0.018,  $df = 7707$ ,  $t = 2.583$ ,  $p < 0.01$ ). The Inconsistent condition affected speakers from a Pharyngealizing background more than Merged listeners ( $\beta = 0.018$ , SE = 0.007,  $df = 7645$ ,  $t = 2.71$ ,  $p < 0.007$ ); here, there was no significant difference between Pharyngealizers and Pharyngeal listeners.

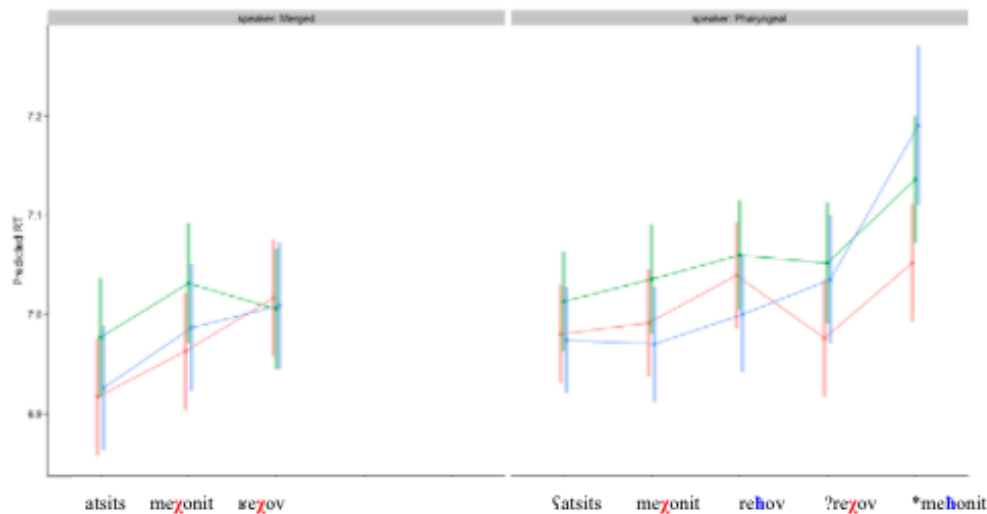
Another interaction between Group and Condition was that Merged listeners responded more slowly to the /h/ condition compared with listeners from pharyngealizing backgrounds ( $\beta = -0.011$ , SE = 0.004,  $df = 7633$ ,  $t = -2.501$ ,  $p < 0.012$ ); a triple interaction between Voice, /h/ and Group, suggests that this effect was mainly due to consistent productions of [h] (i.e. items produced in the Pharyngeal voice;  $\beta = 0.01$ , SE = 0.008,  $df = 7621$ ,  $t = 2.283$ ,  $p < 0.03$ ).<sup>6</sup> That is, while listeners from Pharyngealizing backgrounds responded to consistent /h/ production faster compared with Inconsistent productions, the opposite was true with Merged listeners.

Age and Running Mode were both marginally significant, such that older participants were slower ( $\beta = 0.002$ , SE = 0.001,  $df = 74.9$ ,  $t = 1.95$ ,  $p = 0.054$ ), as were online participants ( $\beta = 0.056$ , SE = 0.029,  $df = 75.03$ ,  $t = 1.95$ ,  $p < 0.055$ ). Figure 3.3 is a visual summary of the model.

### 3.3.3 Discussion

The results revealed a significant difference in the patterns of reaction to switched-category items between Pharyngealizers, Pharyngeal listeners and Merged listeners. Particularly, participants from the first two groups rejected Pathological items ( $\chi \rightarrow \text{h}$ ), or responded to these items significantly more slowly, indicating a difference in the status of this phonetic category in their lexicon.

<sup>6</sup> Recall that in the Merged speaker's voice, /h/ surfaces as [ $\chi$ ].



**Figure 3.3:** A summary of the linear regression model for RTs in Experiment 3. **Blue:** Pharyngealizers, **Green:** Pharyngeal listeners, **Red:** Merged listeners. **Left panel:** Merged voice, **right:** Pharyngealizing voice. **From left to right:** Baseline, /χ/, /ħ/, Dialect-Inconsistent [χ], Pathological [ħ].

Pharyngealizers and Pharyngeal listeners additionally rejected and responded more slowly to Inconsistent items ( $\hbar \rightarrow \chi$  in the Pharyngealizing voice) at a higher rate than Merged Listeners, indicating that the perceived identity of the speaker affected their lexical decision as well.

Merged listeners were slower than the other groups in lexical decision for consistent [ħ] items produced in the Pharyngeal voice. This can be an ‘accent effect’, i.e. the price of processing a variant that is not often met by these speakers. On the other hand, Pharyngealizers were slowed by items produced with [χ] significantly more than Pharyngeal Listeners, suggesting a similar price in the opposite direction. Note that this effect was obtained in the Merged voice, that is, in addition to the effect of slower RTs in the Inconsistent condition.

Overall, the results confirm that some speakers who produce a Merged [χ] category still maintain a covert lexical distinction between [ħ] and [χ]: the Pharyngeal Listeners group. With respect to the Merged listeners group, our interpretation of the results is that the lexical representations of Merged listeners does not include the category [ħ], which is therefore always mapped into an underlying /χ/. Under this interpretation, Merged listeners can perceive [ħ] as a distinct sound, but there is no category in their lexicon that matches it – [ħ] is perceived as a surface variant of /χ/, that has social, but not lexical consequences.

These results are in accordance with the prediction of the “fuzzy representations” hypothesis presented in Darcy et al. (2013): [ħ] is perceived as distinct phonetically, but is assimilated at the phonological encoding level to [χ]. A crucial difference between the current case and the L2 cases that Darcy et al. set out to explain is the distribution of the categories in the input listeners are exposed to. For L2 learners, the perceptual merger (of [o] and [ö], for example) is due to

the absence of a category in their L1. The input they receive in L2 is generally *not merged*.<sup>7</sup> In the context of a native merger, listeners are exposed to merged and non-merged speech at varying degrees.

This brings us to another important difference between the cases: there is always a social context for a native merger, such that one variant is likely to become socially marked, i.e., associated with a particular group of speakers. Thus, *socially*, from the perspective of a Merged speaker, only Mizrahi speakers may produce [h]. On the other hand, *lexically*, only some lexical items that include [χ] may be produced with [h], making generalization more difficult, particularly when exposure to Pharyngealizing speakers is limited in quantity. Therefore, it is easier to attribute [h] to a social structure rather than a lexical distinction.

### 3.4 General discussion and conclusions

Modern Hebrew has often been described as having two major dialects: one in which pharyngeal and non-pharyngeal phonemes are distinct, and another in which pharyngeals are lost altogether. This suggests that a speaker of the merged variety, who lacks pharyngeals in production, would also lack any lexical representation of the distinction between pharyngeal and non-pharyngeal segments, and may have trouble distinguishing the two in perception (as, presumably, would an L2 learner of Hebrew). While such speakers may indeed exist, our results show no evidence of them. Rather, we see two different patterns of speakers who produce the merged category, highlighting the intricacy of the actual linguistic situation that is the outcome of a socially salient merger.

The Merged Listeners in our sample do not behave like “naive” listeners, and are shown to be quite adept at distinguishing between [h] and [χ] – although this phonetic perception is translated only into social, and not lexical, evaluation. As for the Pharyngeal Listeners in our sample, their patterning with the Pharyngealizers is evidence that they too – or at least, some of them – have separate lexical representations for [h] and [χ], despite the identical surface realization. This result is in line with Sumner and Samuel (2009), who demonstrated that rhotic speakers from the NYC area that grew up exposed to the non-rhotic NYC variety are primed by non-rhotic stimuli in a fashion similar to non-rhotic speakers, as opposed to rhotic speakers without such exposure. Crucially, our results, like that of Sumner and Samuel, are in the context of an ongoing change in progress (the adoption of rhoticity in the case of NY). Indeed, it may be the case that such a discrepancy between dialect production and dialect perception and representation is an ephemeral stage typical only of such contexts, since the next generation of speakers would have no phonetic evidence from which to form the divergent representation.

The current case study joins a growing body of work proposing that speech processing proceeds in two parallel routes: indexical (social) and lexical (Sumner et al. 2014; Cai et al. 2017). In

---

<sup>7</sup>This is perhaps not the case in some contexts of second language learning, since learners are often exposed to merged speech produced by their peers.

the rest of the discussion, we address the model proposed by Cai et al. 2017.

The model suggests (although Cai et al. 2017 do not state so explicitly) that during acquisition, we learn what acoustic information is relevant for lexical distinctions and what information is relevant for social structures. In an ideal case, the listener is experienced with their language through a variety of speakers from various backgrounds. Data from the acoustic stream can therefore be more easily divided into lexically-relevant variation (i.e., the phoneme inventory of the language) and socially-relevant variation (individual or group level, e.g., accent). For L2 learners, separating the social and lexical information is not a trivial task, since they are typically exposed to less of the language in general, and more crucially to fewer speakers and accents. It seems plausible, then, that L2 learners may confound lexically relevant dimensions of acoustic variation with social ones.

In our case study, Merged listeners seem to attribute the use of [h] to a relevant social distinction, and can identify it reliably (Experiment 2), but cannot use this variable lexically (Experiment 3), although they are aware of its lexical status through direct instruction in schools and the orthographic system. Pharyngealizers and Pharyngeal listeners are equipped to make the lexical distinction, as demonstrated by their rejection of non-words in which [χ] had been replaced with [h] (the Pathological condition in Experiment 3).

Moreover, for some in participants in the Pharyngealizers and Pharyngeal Listeners groups,  $h \rightarrow \chi$  ('Dialect-Inconsistent') items were also unacceptable. These items were similarly responded to significantly more slowly by these groups, contrary to the Merged listeners groups. Under a dual-route model, this is expected since the Pharyngealizing speaker is expected to have accurate lexical representations of [h] and [χ], and produce them as distinct. This finding is another illustration of the kind of reasoning processes involved in speech processing in the indexical route (Cai et al. 2017).

The last insight, again relevant mostly for the Merged listeners group, has to do with orthography. Most Merged listeners did not reject "Pathological" [h] items despite the fact that the "Pathological" condition is inconsistent not only with the lexicon of Pharyngealizers, but also with the Hebrew writing system (as discussed in Section 3.1.1). That is, most literate Hebrew speakers, most of the time, did not seem to care that the Pathological forms fail to reflect the orthographic reality of the language. Orthographic representations were previously shown to affect online processing of auditory stimuli. This sensitivity to orthography may be limited based on the type of task, as well as the type of phoneme-to-grapheme correspondence. Both interpretations are possible here: first, in most previous studies the task was more meta-linguistic, e.g., rhyme detection (Chéreau et al. 2007; Grainger and Ziegler 2011; Ziegler et al. 2004), while the current task was lexical decision. Meta-linguistic tasks may incline participants to use various heuristics, including their orthographic knowledge, particularly when orthography is likely to be helpful, as with rhyming. Such heuristics may be less active in a word recognition task. Second, the correspondence between graphemes and phonemes was usually more complex in previous studies compared with the current case (for example, *lane* and *rain*). The letters representing [h] and [χ] are absolutely interchangeable in terms of their production in all words used in Experiment 3.

The case of the Modern Hebrew [h]~[χ] merger stands out in two additional ways. First, it is a case of consonants merger, while the majority of the mergers studied so far involve vowel categories (see Section 3.1.2 and references there). There is reason to believe that consonants and vowels play different roles in lexical encoding (Nespor et al. 2003) and speech processing (Turnbull and Peperkamp 2017), making this data point particularly valuable. Second, the non-merged variety is at the same time less prestigious and the one consistent with the orthographic system, as well as with unique phonological alternations for each category. Orthography and prestige go hand in hand in many cases (e.g. Sumner and Samuel 2009). The fact that [h] is not retained in the production of most second generation listeners suggests that orthographic and phonological transparency is inconsequential to language change compared with the role of prestige.

## IV

### Exploring phonological variation with the IAT: Advantages and limitations

Recent studies on online phonological processing have confirmed the intuition of probably any speaker dealing with linguistic variation: listeners have an advantage in processing vernaculars familiar to them. This is reflected in higher accuracy rates in comprehension in noisy environments (Clopper and Bradlow 2008; Adank et al. 2009) and faster reactions in lexical decision tasks (Flocchia et al. 2006; Sumner and Samuel 2009). How this growing expertise is achieved, how it is represented and whether it is triggered given contextual cues, or simply available to the savvy listener at any time – these are still very much open questions. The current study aims to broaden our understanding of two aspects of experience-based efficiency in dealing with phonological variation: its context-sensitivity, and its degree of automatism. The experiments are based on naturally occurring cases of co-variation in Hebrew, in which a marked variant is associated with at least two distinct and salient social groups, that differ in their production of a variety of other segments.

The chapter explores two case studies. The first case, sketched in Table 4.1, illustrates the interaction between two marked variants: the production of [h] (unmarked variant: a uvular trill/fricative [χ]), and the production of [ej] (unmarked variant: [e]). [h] is characteristic of at least two distinct personae of native Hebrew speakers: Old Mizrahi and Broadcaster, who do not merge this category with [χ], which is common in most native speakers (see chapter 3). [ej] is characteristic of Broadcaster-style speech; merging this category with [e] is characteristic of Mizrahi speakers. While these groups may overlap, they are associated with different stereotypes and differ in prosody. The personae presented in the experiment are marked in blue in Table 4.1.

		[χ]~[h]	
		Yes	No
[e]~[ej]	Yes	Broadcaster	Old Ashkenazi
	No	Old Mizrahi	Unmarked

**Table 4.1: The distribution of [e]~[ej] and [χ]~[h] mergers in the natural speech of each guise in Experiment 4 (marked variant in boldface). In order to predict the feature value of one variable based on another, one has to rely on assumed speaker identity. The personae presented in the experiment are marked in blue.**

The second case involves non-native variants, that are both marked: the apical trill [r] (unmarked variant = [ʀ]), associated with native Arabic speakers on the one hand and with native Spanish speakers on the other (more specifically in the context of Hebrew, Spanish speakers are for the most part Jewish immigrants from south American, the largest community being from Argentina); and the production of the etymological [ħ] (unmarked variant [χ]), produced as [h] by Arabic speakers and as [x] by Spanish speakers.<sup>1</sup> The Argentinian [x] is typically more front than the unmarked Hebrew category, and never produced as a trill.

Different guises were used to test whether listeners build context-sensitive expectations regarding the co-occurrence of these variants, based on the presumed identity of the speaker. If listeners quickly calibrate their speech processing based on prior experience, we expect it to be easier for them to relate variants based on their distribution in the speech of the “similar” person. Since these distributions are inversely correlated in the different personae, we expect that the guise should affect the direction of association between the marked variants. In other words, we aim to test whether social context can mediate the degree of association between phonetic variants during online processing.

In Section 4.1, I review studies that demonstrate how speech processing might be affected by the speaker’s perceived identity, highlighting open questions in the field. Section 4.2 introduces the Implicit Association task (IAT) paradigm and its successful application in studies of socio-phonetic variation. Sections 4.3 and 4.4 present the experiments. Section 4.5 puts them in the context of auditory linguistic experiments that have used similar methodologies, highlighting the behavioral differences in response to the same stimuli under a different guise, found in Experiment 5. I then outline possible reasons for why the d-score measure (used to assess the results of IATs) did not reveal speaker-conditioned differences in the association between variants, focusing on the difference between the processing of visual vs. auditory stimuli.

#### **4.1 Background: speaker-oriented predictions**

Experienced listeners use knowledge about the speaker in order to better decode speech. In phoneme identification tasks, listeners can use information about the distribution of acoustic cues in different social groups in the categorization process. One example is that if an English speaker is believed to be a woman, an ambiguous sound on the s-f continuum would be more likely to be categorized as [s] than if the speaker is believed to be a man (Strand 1999). Here, perception mirrors the bi-modal distribution of the center of gravity of the strident, which tends to be higher for both phonemes in women (this predictor for center of gravity can be further broken down into two components, body size and gender performance). Recently, D’Onofrio (2018b) showed that phoneme boundaries can shift based on seemingly much more specific social categories – the personae of the “valley girl” and the “business professional”, both of whom stereotypically produce

---

<sup>1</sup> The pharyngeals and [r] are produced by native Hebrew speakers with a Mizrahi inventory, in addition to Arabic speakers (see chapters 1-3). However, young speakers who use these sounds are more likely to be Arab Palestinian; see Section 4.4).



backed variants of the American English vowel [æ]. When participants were told that the person they will be hearing was described as a valley girl/business man, the rate of classification of an ambiguous [æ] ~ [ɑ] sound as [æ] rose significantly, presumably “correcting” for the backed variants expected by these personae, as with the man/woman categories in Strand (1999).

Similar “calibration effects” were readily learned in a very short time in an experimental setting, provided that learners had a systematic – usually social – source to which they could attribute variation. For example, if the speaker seems to be of a different ethnicity than their caregivers, 18-month-old infants accepted their systematic vowel quality divergence (presumably as an accent); they did not accept the same divergence from a speaker who looks more like their primary caregivers (Weatherhead and White 2018). In the same vein, if (and only if) a speaker’s idiosyncratic phonetic alternation proved to be systematic, and could not be attributed to accidental mispronunciations, adult listeners readily adjusted to it (Kraljic and Samuel 2006, Kraljic et al. 2008, Maye et al. 2008).

Adjustment to speaker can also modulate the preferred semantic interpretation of cross-dialect ambiguous words, such as “flat”, which in British English most frequently denotes a single-story residence, contrary to the American English more common adjectival reading. This co-variation of accent and preferred meaning affected results for British speakers, such that an AE-accented speaker induced more American-leaning meanings, but not for American speakers, who are presumably less experienced with British English and may not be aware of the alternative meaning (Cai et al. 2017).

These studies show that associating phonological variation with speaker/ region/ gender-related differences is possible, and probably useful for decoding speech; Strand (1999) and D’Onofrio (2018b) demonstrate that we accumulate knowledge of these associations and use them naturally, and the perceptual learning studies demonstrate that we can acquire them quickly. However, it is unclear to what extent these effects are automatic and general. One can argue that calibration effects, such as the ones described above, are the result of meta-linguistic control: participants perceive the sound they have heard as ambiguous, and adding the assumption “the speaker is a man/Valley girl”, their categorization leans in the expected direction.

Still, there are some aspects of speaker-oriented processing that are clearly automatic. For example, the more one hears an accent, the faster one is to decode it (e.g. Sumner and Samuel 2009; Sumner and Kataoka 2013). Speakers who are accustomed to an accent respond faster to words that include alternations characteristic of that accent, even if they do not produce these alternations themselves. These are the insights we carry into the study of online speaker-oriented processing.

#### **4.2 The Implicit Association Task in context**

The experiments presented in this chapter test the associations between socially ambiguous variants using the Implicit Association Test (IAT; Greenwald et al. 1998). The IAT is a categorization

task, that relies on the idea that if two categories are related in one’s mind, it should be easy to provide the same response when presented with input stimuli that belong to either; if the categories are not related (or negatively related), it would be easier to have a separate response for each. For example, for most humans, insects are associated with negative emotions and flowers with positive emotions, so it should be easier to respond in the same manner to pictures of flowers and to words associated with positive emotions, like *wonderful* or *good*, than to respond in the same manner to pictures of insects and positive words – simply because flowers evoke in most people more pleasant emotions than insects.

In the task, the “ease” of making an association is measured in accuracy rates and reaction times to a classification task, in which stimuli from four categories of two types, objects and attributes (e.g. insects - flowers, Good - Bad) are presented in a mixed order. The participant uses two keys, “left” and “right”, for all the categories. The experiment has three block types: (i) Practice: only one category type is presented (e.g. insects - flowers) and categorized “left” or “right”; (ii) “left” is assigned to two categories of different types, e.g. insects and Good, “right” is assigned to the other two - flowers and Bad; (iii) The attributive categories are reversed, such that “left” is assigned to insects and Bad, and “right” is assigned to flowers and Good (see Table 4.2).

Bases on an optimized calculation of mean reaction times and accuracy rates for block types (ii) and (iii), the subject’s d-score, a number between -1 and 1 that reflects the magnitude of the association between categories, is determined. The average d-score in a community is taken to reflect the strength of association between the relevant category and attribute.

The IAT has proved to be a reliable method for extracting implicit associations, thus avoiding the difficulty related to subjects not feeling comfortable to share their associations explicitly – for example, in cases in which they would reflect bigotry, racism or sexism. In support of the validity of the IAT as an implicit bias detection, Greenwald et al. (2003) point out that in a meta-analysis of 86 independent samples, the average d-score significantly predicted criterion measures, such as judgments, choices, physiological responses, and behaviors.

The task has recently been adapted to the auditory modality, in an attempt to study implicit associations between variants of speech associated with dialect vs. features that are thought to be more idiosyncratic. Campbell-Kibler (2012) showed that speakers associate together features

Block	Trials	Task	Left key	Right key
1	20	Practice: categorization of category I	Insects	Flowers
2	20	Practice: categorization of category II	Good	Bad
3	20	Test: Categorization of I+II	Insects+Good	Flowers+Bad
4	40	Test: Categorization of I+II	Insects+Good	Flowers+Bad
5	40	Reversed practice of Category I	Flowers	Insects
6	20	Test: Categorization of I+II	Insects+Bad	Flowers+Good
7	40	Test: Categorization of I+II	Insects+Bad	Flowers+Good

**Table 4.2: An example of the IAT.**

of the same dialect faster than features that are not necessarily dependent on the dialect. For example, the suffix *-ing* produced with a coronal nasal [ɪn] (contrary to [ɪŋ]) was associated with the monophthong variant of [aɪ] more than with its diphthong variant, implying an association between these two features of South American English (SAE). On the other hand, no association was found between *-ing* and word final t-deletion, for which the source of variation is not dialectal: t retention in American English has been associated, under specific contexts, with education and articulate speech (e.g., Eckert 2008a). This result corresponds nicely with the explicit judgments provided by speakers after the task, which showed a stronger explicit association between [ɪn] and the monophthong. T-deletion was also perceived as a characteristic of SAE in the explicit judgments, but less strongly, perhaps indicating that participants generally associate more prestige and education with non-South dialects.

The results of Campbell-Kibler (2012) indicate that listeners are sensitive to the co-occurrence of variants – or, perhaps more specifically, socio-linguistic markers – in dialects they are familiar with. How are these associations between features represented?

One possibility is that markers, or phonological variants in general, have abstract representations, and these could be bound together in a meta-category based on some organizing principle. This is close to the interpretation given to “traditional” IATs, that test for associations between broad categories (e.g. insects, flowers) and attributes (“good” and “bad”). Under this “meta-category” approach, at least two explanations of Campbell-Kibler’s (2012) results arise.

First, it is possible that [ɪn] was associated with the monophthong based on the meta-category “reduced variants”. There is evidence that English speakers perceive both variants as reduced forms (Campbell-Kibler 2007). In the case of the monophthong variant of [aɪ], this position has phonological support; a monophthong has a simpler nucleus than a diphthong. In the case of [ɪn], its perception as a reduced form is supported by orthographic knowledge, as the deletion of the word final letter *g* is used to signify a change in the place of articulation, from velar to alveolar. Under this interpretation of the association between [ɪn] and the monophthong, it is not clear why t-deletion was not implicitly associated with [ɪn] as well: t-deletion is equally marked, and can equally be viewed as reduction.

The second possible interpretation within the “meta-category” approach is that participants relied on the co-occurrence of these variants in the speech of SAE, binding these variants into the meta-category “markers of South American English”. This is the explanation preferred by Campbell-Kibler (2012), as it also explains the lack of association between [ɪn] and t-deletion. Under this interpretation, “reduced” and stereotyped variants may be associated with unmarked variants within the same dialect.

A third option, not explored so far, is that listeners based their association relying on *the persona of the speaker* – in this case, a native SAE speaker – as the organizing principle. I borrow the term “persona” from D’Onofrio (2018b), to denote some recognizable speech style that is not necessarily only an accent, but a character packed with other characteristics. For example, “the businessman” or “valley girl” personae in D’Onofrio’s (2018b) study share features of speech

that also occur in other specific dialects, but that are strongly associated with these personae. Similarly, the “Broadcaster” persona used in Experiment 4 below is a recognizable speech style that borrows features from various dialects, and is primarily associated with prescriptive speech. Under this interpretation, listeners have detailed characterizations of the distribution of different variants based on the speaker. This proposal is supported by studies that found some variants to be markers when embedded in the speech of one speaker, and at the same time, they could also be of no social significance – and possibly, of different consequences to processing – in the speech of another (Pharao et al. 2014, King and Sumner 2014). For example, Pharao et al. (2014) showed that the fronting of /s/ is perceived as effeminate, or gay, when embedded in the speech of a (male) Standard Danish speaker, but not when embedded in the speech of what the authors refer to as “Street” Copenhagen Danish. In King and Sumner (2014), the authors went further to check not only the salience and social import of markers, but also their effect on online processing. Using a lexical decision task with two voices, they found that words produced with final cluster simplification were processed faster and more accurately when the recorded speaker exhibited AAVE prosodic characteristics, than when she had General American prosodic characteristics, illustrating that processing a marked alternation may become easier when contextual cues for the alternation are available.<sup>2</sup>

A recent study confirms the same intuition, by showing that speakers and listeners alike track the co-occurrence of the alveolar variant of the affix [-ing] across styles and grammatical categories (Vaughn and Kendall 2018). In the first phase of the study, five speakers produced 141 sentences in which [-ing] plays different grammatical roles (e.g. “climbing” vs. “interesting”). Speakers were requested to produce two versions of each sentence: once with the prestigious velar [ɪŋ], and another with the alveolar [ɪn], associated with SAE as mentioned above, but also with more casual, rapid speech. In an acoustic analysis, the authors found that although speakers were not instructed to change any other aspect of their speech, when producing [ɪn] most of them also significantly changed their production of other speech sounds: the diphthong [aɪ] became shorter while front lax vowels became longer, the mid front vowels slightly shifted, and inter-vocalic /t/ was reduced. This co-variation is not random, of course; the vocalic variants are imported from the SAE style, while /t/ reduction characterizes casual speech. It seems, then, that in order to produce [ɪn] speakers tended to perform a full persona, or style.<sup>3</sup>

In a series of subsequent perception tasks, Vaughn and Kendall (2018) chose a subset of the recorded materials, in which co-variation of [ɪn] with the other variables was less significant than in most of the sample. In the first perception task, listeners were requested to rate how accented the speech sample was; it was found that the production of [ɪn] significantly contributed to perceived accentedness. For the next task, [ɪn] productions were masked by white noise. The new

---

<sup>2</sup> In fact, as King and Sumner (2014) mention, the cluster simplification rate in AAVE is not different from other native American English vernacular. Still, it seems that listeners expected more reduction in this variety, perhaps over-generalizing based on other processes.

<sup>3</sup> Note that the production choices of speakers in this study cannot be explained solely based on some principle of reduction. Longer lax vowels, for example, cannot easily be viewed as a type of reduction, yet participants lengthened these vowels significantly, as expected in an attempted performance of an SAE accent.

manipulated stimuli were presented to fresh participants, who were instructed to guess whether the white noise portion masks “in” or “ing”. Participants were significantly above chance level at guessing the correct production; furthermore, the “accentedness” rating of each sentence from the previous task corresponded nicely with predicted [ɪn]. Linguistic constraints also played a role in levels of predicted [ɪn]: listeners predicted significantly more [ɪn] in verbs than in nouns, again demonstrating fine-grained distinctions in perception of structured variation.

The perception study in Vaughn and Kendall (2018) illustrates the roles of social and linguistic knowledge in forming predictions regarding particular variants. The present study aimed to expand on these findings, by exploring two routes:

1. What is the role of awareness/attention in the prediction of co-variation? In Vaughn and Kendall (2018) (as well as in previous studies described in Section 4.1), meta-linguistic control is available to listeners who guess what the masked sound may be, or categorize sounds. The current design involves a fast RT collection paradigm, which was credited for tapping into implicit representations.
2. Vaughn and Kendall (2018) distinguished between two speech styles that are associated with [ɪn], and that otherwise co-vary with different variants (fast vs. slow speech, /t/ reduction vs. vowel shift). Still, the two styles don’t necessarily clash at the persona level: a SAE speaker might go casual and reduce /t/, for example. The same is true for the first case study presented below: there are indeed Yemenite broadcasters. However, in the second case study, the identities are much less likely to be reconciled as one and the same, since both imply that the speaker has a different L1 – Palestinian Arabic or Argentinian Spanish.

#### **4.3 Experiment 4: [ħ]~[χ] and [e]~[ej]**

In Modern Hebrew, producing [ħ] is socially ambiguous, marking (mostly old) Mizrahi speakers on one hand, and elitist media broadcasters on the other (see Chapter 1). Each of these identities comes with its own cluster of other linguistic features. “The broadcaster”, an artificial speaking style produced by old-school radio broadcasters, utilizes the pharyngeals as well as the Ashkenazi phonemic distinction between [e] and [ej]. “The Yemenite”, a sub-category of the Mizrahi ethnic identity, that is widely associated with preserving the Mizrahi pronunciation (Gaftner 2016a), utilizes the pharyngeals but does not produce the [ej]~[e] distinction, producing [e] across the board.

The relevant feature for the current design is that these identities differ in their production of the [e]~[ej] distinction – the Broadcaster retains it, while Mizrahi speakers typically produce [e]. Merging the categories of [ħ] and [χ] into [χ] is the unmarked option in most of the public arenas (see Chapters 2 and 3), and most speakers seem to generally merge [ej]~[e] in production as well (although the diphthong [ej] seems to be preserved by some younger speakers in borrowed

words, as in [imejl] ‘e-mail’). Matras and Schiff (2005), for example, found between 71-97% monophthongs per individual in their sample, that included natural recordings from a variety of speakers.

The goal of the experiment is to test whether the different guises would affect the online association between markers. If the associations between markers are based on meta-categories that tie between *variants*, it should not matter who the speaker is – the result should yield a similar association. If, however, the associations between markers are grounded in speaker-specific perception (Cai et al. 2017, Montgomery and Moore 2018, D’Onofrio 2018b), we expect to find opposite associations based on the speaker’s perceived identity.

#### 4.3.1 Experiment 4a: “Broadcaster” ([h]+[ej])

**Stimuli:** Five words of each category were selected for the main task (i.e. the IAT). Since [ej] and [h] are phonemic in non-merged varieties, all words used in the categorization task contained one of these phonemes. The words used were: [bejtsa] ‘egg’, [lejtsan] ‘clown’, [mejmad] ‘dimension’, [mejrav] ‘maximum’, [tej] ‘tea’; and [hida] ‘puzzle’, [hibur] ‘essay’, [halil] ‘flute’, [hamor] ‘donkey’, [hag] ‘holiday’.

An 18 years old native Hebrew speaker, native pharyngealizer, was recorded for the Broadcaster guise. All item pairs had exactly the same length of marker offset (e.g. in [hamor]-[χamor], the first phones ended after exactly the same time), to ensure accurate measurement. Examples of the stimuli are available in Table 4.3. The pictures used for the introduction of the speaker were of two different men (see Figure 4.1).

**Procedure:** All the experiments described in this chapter ran on Minno, a platform for running experiments online (Zlotnick et al. 2015). Participants were instructed to wear headphones, and begin the task only when they have 12 minutes of uninterrupted time. They were first presented with the speaker’s picture, along with an explanation regarding the labels used for the categories of the experiment.

**Categorization labels:** The chosen categorization labels require some elaboration, as it has been demonstrated that the mere label of the category may influence d-scores in the IAT (Weirich et al. 2020; a more elaborate discussion of this study is found in the general discussion). The [h]~[χ] distinction is overtly referred to by name – speaking Hebrew with pharyngeals is called

	Exp.4b: Yemenite Guise		Exp. 4a: Broadcaster Guise		
Compatibility	Yes	No	Yes	No	
[h]~[χ]	hamor	xamor	hamor	χamor	‘donkey’
[ej]~[e]	betsa	bejtsa	bejtsa	betsa	‘egg’

**Table 4.3: Examples of stimuli in Experiment 4; each persona exhibits a different alignment of the markers**



**Figure 4.1: The pictures used in the presentation of the speakers. Left: the Broadcaster guise (Experiment 4a); Right: the Yemenite guise (Experiment 4b).**

“ledaber be-Het ve-’ain”, lit: ‘to speak in [h] and [ʕ]’, based on the names of the letters representing the pharyngeals (see Chapter 3).<sup>4</sup> The labels were therefore *be-Het* and *lo be-Het*, ‘in [h]’ and ‘not in [h]’ respectively. The second set of labels was more challenging, since the distinction or merger between [e] and [ej] is less common in discourse; however, relying on a Youtube video that calls the [e] variant “medaber bli Jod”, lit: ‘speaks without Jod’ (“jod” is the name of the letter that represents the vowel [i] and the glide [j] in Hebrew), the chosen labels were *im Jod* “with Jod” and *bli Jod* “without Jod”.<sup>5</sup>

The labels were presented as follows:

There is more than one way to pronounce certain words in Hebrew. In the following experiment, you will hear words in Hebrew in two different versions [recorded by a young radio broadcaster].<sup>6</sup> You’ll need to categorize them based on the category names at the top left and top right corners of the screen.

We will use four categories:

You can produce the words [maʕar] ‘sold’ and [maħar] ‘tomorrow’ in the exact same manner, but you can also produce the latter with a pharyngeal [ħ] (“ledaber be-Het ve-’ain”).

We will call the pharyngeal production *be-[ħ]et*, and the other *lo be-[ħ]et*.

You can produce the word [teʕa] ‘nine’ so that it will sound like [tejʃ], but you can also produce it so that the first vowel would be like in the word [peʕa] ‘crime’. These productions will be referred to as *im Jod* and *bli Jod*, respectively.<sup>7</sup>

<sup>4</sup> *Het* is the name of the letter that represents the voiceless pharyngeal fricative in non-merged dialects; *’Ain* is the name of the letter that represents the voiced pharyngeal approximant.

<sup>5</sup> The video is available at: <https://www.youtube.com/watch?v=51aw0hwNbR8>

<sup>6</sup> The phrase in brackets was only introduced in the Broadcaster guise. In the second experiment, this description was replaced with a short recording of the speaker, in which he presented himself (see Section 4.3.2).

<sup>7</sup> The minimal pair of [h] and [x] was presented orthographically, since the two are represented by different letters, [ח] and [כּ] respectively. The distinction between [teʕ] and [tejʃ] was also presented orthographically, by inserting two ‘jod’ letters, conventionally read as [ej].

These are the four categories of the experiment. Listen and categorize according to the instructions.

The IAT instructions were taken from Project Implicit (Xu et al. 2014), and repeated with minor changes in every intermission between blocks. Block order was randomized across participants, such that approximately half of the participants began with [χ] and [ej] associated with the same key, and the other half with [χ] and [e] associated with the same key. The number of blocks and the number of trials within block, as well as the algorithm for calculating the d-score, are adopted from Greenwald et al. (2003). Table 4.4 outlines the task. Following the task, participants completed a short demographic questionnaire (see discussion in Section 4.5).

**Participants:** Recruitment for the experiments was done online, using social media platforms. The criteria for participation were normal or corrected to normal hearing and no interruptions during the task. Thirty-five native Hebrew speakers participated in the experiment (mean age = 32).

**Results:** A one-sample t-test revealed a medium to strong association between [h] and [ej] ( $t = 3.18$ ,  $Df = 33$ ,  $p < 0.005$ , *Cohen's d* = 0.545), that is, as predicted for the “broadcaster guise”. Block order was not significant, i.e., it did not matter whether the [h]+[ej] association or [h]+[e] association was introduced first ( $t = -0.07$ ,  $Df = 32$ ,  $p > 0.9$ , *Cohen's d* = -0.24).

#### 4.3.2 Experiment 4b: “Yemenite” ([h̄]+[e])

In the “Broadcaster guise”, [h] was found to be associated with [ej]. The following experiment is complementary to 4a: it utilizes the same variables, under a guise in which their association is expected to be inverted. It was designed in order to tease apart between the “Variant meta-category” and “Speaker meta-category” hypotheses, outlined in Section 4.2: if the association found between the variants is different from that obtained in the Broadcaster guise, the ‘Variant meta-category’ interpretation should be abandoned.

Based on the first study, the design of Experiment 4b was improved on two main points. First, in order to lend participants more social and acoustic cues to converge on the “appropriate” persona, a short recording was presented to participants along with the picture of the speaker prior to

Block	Trials	Task	Left key examples	Right key examples
1	20	Baseline [h]~[χ] discrimination	[hibur]	[χibur]
2	20	Baseline [e]~[ej] discrimination	[mejmad]	[memad]
3	20	Test: Categorization of [h]+[ej]	[hibur]+[mejmad]	[χibur]+[memad]
4	40	Test: Categorization of [h]+[ej]	[hibur]+[mejmad]	[χibur]+[memad]
5	40	Reversed practice: [h]~[χ] discrimination	[χibur]	[hibur]
6	20	Test: Categorization of [h]+[e]	[hibur]+[memad]	[χibur]+[mejmad]
7	40	Test: Categorization of [h]+[e]	[hibur]+[memad]	[χibur]+[mejmad]

**Table 4.4: Block order and examples of stimuli in Experiments 4a-b.**



the classification task. When the picture is clicked, the speaker presents himself by name, age and place of residence, that was chosen for being known as a majority-Yemenite town (Rosh ha-'Ain). In this short self-presentation, the speaker did not use words that included the target variables of the classification task. Participants could listen to this recording for as many times as they wanted.

Second, two questions were added to the post-task questionnaire, targeting participants' overt views regarding the relationship between the variables:

1. A multiple choice question: "which of the variants are related in your opinion?" with the possible answers: "[h] and [ej]", "[h] and [e]", "the variants are not related".
2. An open question: "Why are these variants related in your opinion?"

The overt views of participants serve both a sanity check for the guise (are participants aware of the association of the particular variants with the persona of the speaker?), and as a test of whether overt and covert views regarding the variables are compatible (as found in a previous auditory IAT; Campbell-Kibler 2012).

**Stimuli:** The same words used in Experiment 4a were recorded for the Yemenite guise by a 60 years old man, native pharyngealizer. The categorization labels and procedure were identical to those used in Experiment 4a.

**Participants:** Seventy-four native Hebrew speakers participated in the experiment (mean age = 40.5). The recruitment method and criteria for participation are as in Experiment 4a.

**Results:** A one-sample t-test of the d-scores revealed a medium to strong association between [h] and [ej] ( $t = 3.8$ ,  $Df = 73$ ,  $p < 0.001$ , *Cohen's d* = 0.44), contrary to the prediction for the "Yemenite guise". Block order was not significant, i.e. it did not matter whether the [h]+[ej] association or [h]+[e] association was introduced first ( $t = -1.46$ ,  $Df = 72$ ,  $p = 0.15$ , *Cohen's d* = -0.34).

Overtly, most participants (36) agreed that [e] and [h] are related. Seven participants claimed the opposite association and 31 found the variables to be unrelated. The fact that roughly half of the participants admitted the [e]+[h] association runs against the results obtained in the implicit task (the IAT). Furthermore, there were no statistically significant differences in d-scores between participants based on their overt judgments, as determined by one-way ANOVA ( $F(2,71)$ ,  $p = 0.4$ ).

#### 4.3.3 Success rates in categorization of target sounds in Experiments 4a+b

A merged set of data was constructed from the results of studies 4a and 4b, in order to assess the general success rates in categorization, based on the first two practice categorization blocks (blocks 1,2). In block 1, participants had to categorize [h] and [χ], and in block 2, they categorized [e] and [ej]. Therefore, it was not possible to compare the success rates in all four stimulus types within the same model. In order to compare the relative success rates, the following models were constructed:

- To assess whether the distinction between [e] and [ej] was easier to make than the distinction between [h] and [χ] (or vice versa), a model with the fixed effects of Categorization Block ([e]-[ej] or [h]-[χ]) and Guise (Broadcaster/Yemenite) was used, with the random intercepts Participant and Item.
- To assess whether one of the categories was more easily identified (e.g., whether participants responded to [ej] faster or more accurately than to the other category [e]), a model with a subset of the data (only block 1) was used, with the fixed effect Stimulus Type and Guise (Broadcaster/Yemenite), and the random intercepts Participant and Item. A model for block 2 was constructed with the same question in mind.
- For each type of analysis, a generalized linear regression was used for accuracy and a linear mixed effects regression for the (log-transformed) RTs.

**Block effects:** There was no significant difference in accuracy ( $\beta = 0.44$ ,  $SE = 0.28$ ,  $z = 1.598$ ,  $p > 0.1$ ) or RTs ( $\beta = 0.02$ ,  $SE = 0.05$ ,  $Df = 34.13$ ,  $t = 1.48$ ,  $p > 0.1$ ) between the two blocks, i.e. participants were as successful in categorizing [e] and [ej] as they were with [h] and [χ]. However, the Guise significantly affected categorization: participants were more successful in the Broadcaster compared to the Yemenite guise (Accuracy:  $\beta = 0.4$ ,  $SE = 0.13$ ,  $z = 3.15$ ,  $p < 0.005$ ; RTs:  $\beta = 0.13$ ,  $SE = 0.03$ ,  $Df = 4.08$ ,  $t = 4.084$ ,  $p < 0.0001$ ). There was also a marginal interaction between Guise and Block in the RT model, such that items in the [h]-[χ] categorization were responded to more slowly in the Yemenite guise ( $\beta = 0.05$ ,  $SE = 0.03$ ,  $Df = 107$ ,  $t = 1.8$ ,  $p = 0.075$ ).<sup>8</sup> The likely cause of this difference is that the recordings of the stimuli were different in each guise.

**Stimulus effects:** The marked category [h] was identified faster than its non-marked counterpart [χ] in both guises ( $\beta = 0.11$ ,  $SE = 0.04$ ,  $Df = 9.1$ ,  $t = 2.9$ ,  $p < 0.02$ ). It was also responded to marginally more accurately ( $\beta = 0.44$ ,  $SE = 0.24$ ,  $z = 1.86$ ,  $p = 0.06$ ).<sup>9</sup> The guise also significantly affected categorization of [h]~[χ], again with marginally higher accuracy and significantly faster identification in the Broadcaster guise (Accuracy:  $\beta = 0.28$ ,  $SE = 0.15$ ,  $z = 1.9$ ,  $p < 0.06$ ; RTs:  $\beta = 0.18$ ,  $SE = 0.04$ ,  $Df = 129.8$ ,  $t = 3.88$ ,  $p < 0.001$ ).

The marked category [ej] was identified faster than its non-marked counterpart [e] in the Broadcaster guise, but not in the Yemenite guise (main effect of [ej]:  $\beta = -0.08$ ,  $SE = 0.05$ ,  $Df = 8.38$ ,  $t = -1.71$ ,  $p < 0.12$ ; interaction of [ej]\*Broadcaster:  $\beta = -0.04$ ,  $SE = 0.02$ ,  $Df = 2061$ ,  $t = -2.15$ ,  $p = 0.031$ ). There was no effect of stimulus type on accuracy. The same main effect of Guise, found in the previous models, was found here as well (Accuracy:  $\beta = 0.49$ ,  $SE = 0.15$ ,  $z = 3.307$ ,  $p < 0.001$ ; RTs:  $\beta = 0.15$ ,  $SE = 0.03$ ,  $Df = 129.6$ ,  $t = 4.544$ ,  $p < 0.0001$ ).

Table 4.5 summarizes the accuracy rates and mean reaction times in the practice categorization blocks used for the models described above, by stimulus and guise.

<sup>8</sup> The full models can be found in Appendix C.

<sup>9</sup> These results are congruent with the results of the categorization task of [h] and [χ] presented in Chapter 3.

stimulus	Error rate		Mean RT in ms (SD)	
	Broadcaster	Yemenite	Broadcaster	Yemenite
[ej]	6%	11%	1341 (418)	1679 (433)
[e]	8%	18%	1450 (483)	1897 (595)
[h]	9%	15%	1213 (368)	1679 (667)
[ç]	11%	22%	1265 (340)	1728 (679)

**Table 4.5: Mean error rates and reaction times by Stimulus Type and Guise in the practice blocks. RTs of above 2.5 standard deviations the participant’s mean or below 150ms were discarded.**

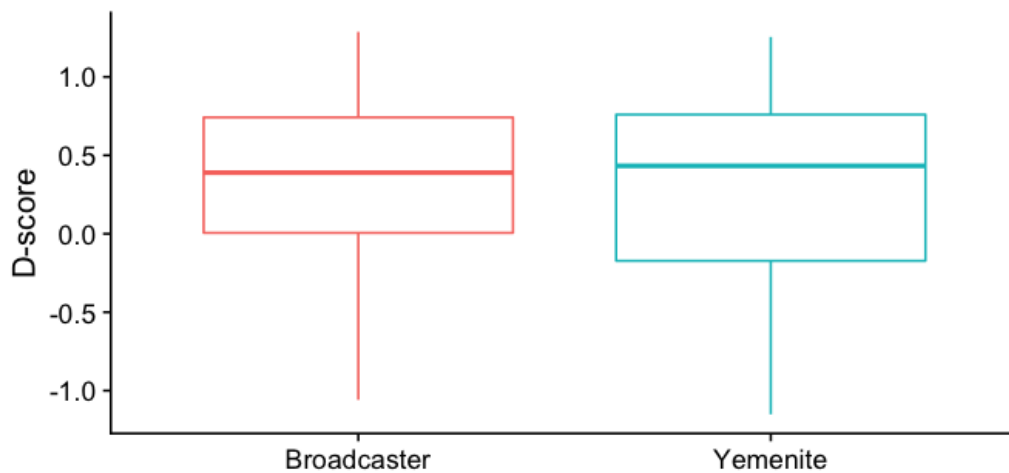
Association	Stimulus	Error rate		Mean RT in ms (SD)	
		Broadcaster	Yemenite	Broadcaster	Yemenite
[ej]+ [h]	[ej]	6.7%	7.5%	1416 (718)	1600 (709)
	[e]	8%	10.8%	1469 (715)	1755 (772)
	[h]	4.2%	12%	1208 (622)	1627 (818)
	[ç]	5.4%	13.4%	1452 (752)	1702 (827)
[e]+ [h]	[ej]	5%	7.2%	1541 (717)	1717 (744)
	[e]	11.8%	12%	1659 (826)	1900 (845)
	[h]	6.7%	11%	1343 (626)	1776 (846)
	[ç]	12%	20%	1714 (909)	1943 (921)

**Table 4.6: Mean error rates and reaction times by guise, stimulus and association in target blocks, Experiment 4. RTs of above 2.5 standard deviations above the participant’s mean or below 150ms were discarded.**

Table 4.6 summarizes the results of the target blocks: reaction times and error rates for each association. Overall, the RTs and error rates were of the same order of magnitude found in previous auditory IATs (e.g. 12.2% errors in the categorization of the velar vs. alveolar variant of the English *-ing* in Campbell-Kibler 2012; 16% for categorization of the German [ç] vs. [ç] variants in Weirich et al. 2020; and 1196-1475ms on average for the auditory stimuli in Weirich et al. 2020).

#### 4.3.4 Discussion

Experiment 4 revealed a medium to strong association between [ej] and [h]. This was expected under the Broadcaster guise, but not under the Yemenite guise. Figure 4.2 illustrates the distribution of d-scores according to guise. As a Welch test confirmed, there was no significant difference between the two guises ( $t = 0.28656$ ,  $df = 69.774$ ,  $p > 0.7$ ). Furthermore, a point-biserial correlation run on the overt judgements of participants in the Yemenite guise (the expected association “[e]+[h]” was coded as 1, the unexpected associations “no association” and “[ej]+[h]” were coded as 0) found that the d-score was not correlated with the overt judgements of participants ( $Df = 72$ ,  $t = -0.87458$ ,  $p > 0.3$ ). This result is incongruous with previous auditory IATs, that found the overt judgements of participants to correlate with their d-scores (Campbell-Kibler 2012).



**Figure 4.2: Distributions of D-scores as a function of speakers' guise in Experiment 4.**

The results are compatible with at least two interpretations, depending on the meta-category that listeners are allegedly relying on for categorization:

1. The “marked/non-reduced variants” interpretation: [h] and [ej] are bound into the meta-category of “socially marked variants”, regardless of the identity of the speaker. It is therefore easier for listeners to group them together, as the contrast to the common variants [χ] and [e]. Under this account, any two marked variants should be grouped together more easily compared with unmarked variants. This account falls short of explaining why in Campbell-Kibler (2012), an association was found between [ɪn] and the monophthong variant of [aɪ], but not between [ɪn] and t-deletion.
2. The “persona identification failure” interpretation: context did affect the results; listeners simply interpreted both guises as “the unmarked speaker”, who uses [χ] and [e]. Under this account, the persona of the speaker was not emphasized enough, and participants relied on their familiarity with “the average speaker”. This interpretation is undermined by the results in the “Yemenite” guise; the overt judgements of participants under this guise, which lean towards an associate between [h] and [e], indicate that participants acknowledge this marker grouping. More specifically, the fact that no correlation was found between the overt judgement and the implicit task, seems to be at odds with the assumption that the identity of the speaker and its linguistic correlates were not available to participants.

The dominance of the [h] and [ej] association could have been strengthened by differences in the social status of the variants used in the experiment: while the [χ] and [e] variants are unmarked, the [h] and [ej] variants are marked, even stereotyped in the case of [h]. This could have biased

participants towards a categorization that is based on the social status of the variants (“marked variants”), rather than relying on their persona-conditioned correlations, that, in principle, could be fleshed out under different circumstances. In other words, the social status of the variant itself might be more salient than any persona, causing participants to rely on the global status of the variants rather than relying on contextual cues.

In order to resolve this confound, the target variants in the task should all be of the same social status, i.e., markers. This is achieved in the next case study.

The design of Experiment 5 also addresses a more general weakness of the original IAT – the inherent ambiguity of the results, noted in Greenwald and Farnham (2000): “Because it uses complementary pairs of concepts and attributes, the IAT is limited to measuring *the relative strengths of pairs of associations* rather than absolute strengths of single associations” (p. 1023, my emphasis). In the case of Experiment 4, it is unclear whether the underlying association which derives the results is “[χ]+[e]” or “[h]+[ej]”, and either case bears different theoretical implications as outlined above.

Karpinski and Steinman (2006) designed a modified version of the IAT, that addresses this problem: the Single Category IAT (hereafter SC-IAT). In the SC-IAT, there are two “attribute” categories (GOOD/BAD) but only one object (e.g. Dogs). In the current case, the “attribute” categories are translated into the non-ambiguous variants, i.e. the variants that are unambiguously associated with one persona but not the other; the “object” is the ambiguous variant, associated with both personae.

#### **4.4 Experiment 5: [h]~[x] and [r]**

Experiment 5 focuses on an L2 ambiguity of the apical trill/tap variant of the Hebrew rhotic. The trill is associated with some groups of native Hebrew speakers (mostly Mizrahi, but also “the Broadcaster”); however, when produced by young speakers, it is associated primarily with a variety of non-native speakers, among whom L1 Arabic speakers – Palestinians; and L1 Spanish speakers – Jewish-Argentinian immigrants, which are the personae introduced in the current design. Most native speakers produce the Hebrew rhotic as [ɾ].

In addition to [r], the other relevant feature of these guises is that they differ in their production of the etymological [h], produced by the majority of young native Hebrew speakers as [χ]. In Arabic, it is produced as [h], while Spanish speakers typically produce it as [x]. This is one of the variables used in Experiment 4, but note that on the current case, both options are socially marked (i.e. not variants produced by the majority of Hebrew speakers; see Table 4.7).

These personae were chosen based on their presumed salience for Hebrew speakers: Argentinian immigrants and Palestinians are represented, socially and linguistically. The two identities were also chosen due to their compatibility with a *young* persona. Unlike the production of the pharyngeals, which is attested among young native speakers in some Mizrahi communities, it is

	Unmarked variant = [χ]	
	[ħ]	[x]
[r] (unmarked variant = [ʁ])	L1 Arabic	L1 Spanish

**Table 4.7: The alignment of features according to guise in Experiment 5.**

very rare to find young native Hebrew speakers who produce [r]; as mentioned above, most young native speakers would typically produce [ʁ]. On the other hand, many Palestinians and Argentinians use this variant when speaking Hebrew, relying on their native language. Since Argentinians of various ages continue to immigrate to Israel, the marker is still very much alive for younger speakers as well.

The current choice of variants and personae allows us to control for a multitude of factors:

1. All three categories used in the current experiment are socially marked. [r] and [ħ] are rarely produced by native Hebrew speakers; the young generation of speakers generally produces [ʁ] and [χ] for the corresponding phonemes. Even when produced by native speakers, these variants are marked. [x] appears only in the speech of non-native speakers, including L1 Spanish speakers.<sup>10</sup>
2. On a related note, the current case is also, in a sense, simpler than the one used in experiment 4, because it does not directly involve a merger: [ħ] and [x] are both marked variants, the unmarked counterpart of both is [χ].
3. In the previous set of experiments, overt judgements ("Are the features related? why?") were only collected in the Yemenite guise. The current design allows us to test whether the guise primed speakers to expect a particular association between the features more often. In addition, it allows us to explore the correlation between overt and covert associations.

**Stimuli:** The stimuli for both guises were recorded by a single native Hebrew speaker, a 40 years old trained phonetician, who is also a native Spanish speaker with an Argentinian family background, with an intermediate knowledge of Palestinian Arabic (speaking and writing). Since Arabic is not a native language of the speaker, the recordings were sent to three native Arabic speakers in order to ensure naturalness.

Five words of each category were selected for the main task (i.e. the SC-IAT). For /h/, the words are [maħit] ‘mash’, [meħake] ‘wait’, [meħika] ‘erasure’, [meħila] ‘tunnel’, [meħasen] ‘vaccinate’. The words were recorded once with [ħ] (Arabic variant) and once with [x] (Argentinian

<sup>10</sup> Despite the greater familiarity in the general public with the L1 Russian accent, which also typically includes this variant (1 million immigrants from the USSR, most of them Russian speakers, have arrived to the state in the 1990s), the Argentinian guise was chosen in order to include the ambiguous marker [r], that is not shared with L1 Russian.

variant). The [h~x] words were chosen such that none of them included environments in which other markers should appear, that would give away the identity of the speaker. For example, voiced intervocalic stops, that undergo spirantization in Spanish, were not included. This was done in order to ensure the relative neutrality of words that are being categorized, because the same recordings of [h] and [x] words were used in both guises (contrary to the previous design, which, due to the typically different ages of the used personae, was recorded separately for each guise).

On the other hand, in the [r] items (the association category, i.e. the category that is only measured through its relation to [h~x]), many characteristics of an Argentinian Spanish/Palestinian Arabic pronunciation were included. This was done in order to strengthen the resolution of the speaker's identity. The words are: [margiʕa] 'calming', [merageʃ] 'exciting', [marʕiʃ] 'noisy', [merageʃ] 'spy', [merkaz] 'center'. The three markers (other than the target variant) that were used were: devoicing of [z] in Spanish (in the word [merkaz]), spirantization of [g] in Spanish (in the words [marʕiʃa], [merageʃ], [merageʃ], [marʕiʃa]) and the production of the voiced pharyngeal approximant in Arabic ([margiʕa], [marʕiʃ]). Again, the onset of the target sound was identical for stimuli of all three categories.

The pictures used for the introduction of the speaker were of two different men, presented as "Waseem" (Arab) and "Manuel" (Argentinian). Each guise was presented by a short recording of the same speaker who produced the stimuli for the SC-IAT. The sentences, like the [r] words, included many linguistic indicators to the identity of the speaker (marked in blue below). The included markers were: spirantization and [h]-deletion in Spanish; long vowels, glottal and pharyngeal productions in Arabic. None of the target sounds for the categorization task were included in the self-introduction segment. The first sentence was produced in the L1 of the guise, and immediately translated to Hebrew.

- Argentinian guise:

ola! jo soj manuel - ani manuel, asiti alija **β**e-ani kan ba-me**ð**ina fmone janim. ani iklateti et a-milim la-nisuj je-mijad tiʃtatfu **β**o.

'Hola! Yo soy Manuel. I'm Manuel, I made "Aliya" and I have been in the state for eight years. I recorded the words for the experiment you are about to participate in.'<sup>11</sup>

- Arab guise:

is-sala:m ʕalejkum, ʔismi wasi:m. ʔani wasi:m, ʔani lomel min**h**al ʕasakim ve-kalkala ba-ʔakademit telʔaviv jafo. ʔani **h**iklateti et **h**a-milim la-nisuj je-mijad tiʃtatfu bo.

'Hi, I'm Waseem, I grew up in Taybeh and I'm studying business administration and economy at the academic college Tel-Aviv Yafo. I recorded the words for the experiment you are about to participate in.'

---

<sup>11</sup> *Aliya*, lit. "ascension", is the name given to Jewish immigration in Hebrew.

**Procedure:** The design was adopted from Karpinski and Steinman (2006). The SC-IAT is identical to the IAT, with two exceptions: (i) there is no practice block for the single variant, in this case [r]; (ii) there is no “switch block” in which participants only practice at categorizing the switched key for the single variant. The order of the blocks is presented in Table 4.8. Each participant was randomly assigned a guise, such that half of the participants were exposed to the “Palestinian” condition and half to the “Argentinian” condition.

**Categorization labels:** While the variant [r] has a name in Hebrew, *resh mitgalgelet* lit. ‘rolling r’, as does the pharyngeal production [h] (*ledaber be-Het ve-’ain*, see Section 4.3), there is no agreed-upon name for the [x] variant. I named it *Het raka*, lit. ‘soft Het’, since it is less intense than the common uvular trill/fricative production. Sample recordings of the target variants between two vowels were introduced along with the categorization labels. These recordings were produced by a different speaker and presented in both guises. After the introduction of the labels and their example reference, the picture of “Waseem”/“Manuel” was introduced, along with the speaker’s short self-presentation.

Following the SC-IAT task, participants completed a short demographic questionnaire, which included questions about their intuitions regarding the overt association between the variants. As a “sanity check”, three questions were added:

1. Where is the speaker who recorded the stimuli originally from? [This question was included to ensure that participants listened to the instructions presenting the guise]
2. What social groups do you know who use the variant? [This question was asked regarding each of the markers categorized in the experiment]
3. What other features of an L1 Arabic/Argentinian Spanish accent do you know? [This question was meant to assess the overall familiarity of the accent in the population]

#### 4.4.1 Experiment 5a: “Palestinian”

**Participants:** Forty-two native Hebrew speakers were randomly assigned to the “Palestinian” guise (mean age = 25.7). The recruitment method and criteria for participation are as in the previous experiments.

Block	Trials	Task	Left key examples	Right key examples
1	20	Baseline [h]~[x] discrimination	[mehika]	[mexika]
2	20	Test: [h]+[r] association	[mehika], [meragel]	[mexika]
3	40	Test: [h]+[r] association	[mehika], [meragel]	[mexika]
4	20	Test: [x]+[r] association	[mehika]	[mexika], [meragel]
5	40	Test: [x]+[r] association	[mehika]	[mexika], [meragel]

**Table 4.8: The procedure and sample stimuli of the SC-IAT in Experiment 5.**



**Results:** Participants' error rates, particularly with [h~x] items, were much higher than in the previous experiments (mean error rate for [x] = 24.7%, mean error rate for [h] = 25%). [h] and [x] seem to be less distinguishable at the acoustic level compared with the contrasts used in Experiment 4 ([χ] vs. [h] and [e] vs. [ej]). The norm for calculating the d-score of a given participant is a cutoff point of 25% in one of the four target blocks. The traditional cutoff had left only 18 participants for the analysis. Therefore, a more permissive cutoff point of 25% errors *overall* in the categorization of [h~x] was chosen. This cutoff left 25 participants for the analysis, with a mean error rate of 14% for [x] and 16% for [h]. A more detailed analysis of errors under both guises is presented below, after the results of 5b.

The analysis of d-scores yielded no significant association in either direction ( $t = -0.49, Df = 17, p = 0.63$ ). Block order had no effect either ( $t = -0.25, Df = 15.92, p = 0.8$ ). This was the case with and without the exclusion of participants with high error rates.

The overt judgements of speakers confirmed that most participants supported the expected association under the Palestinian guise: 21 associated [h] with [r], 15 participants marked [x] and [r] as related, and 7 found no relationship between the variants.

#### 4.4.2 Experiment 5b: "Argentinian"

**Participants:** Forty-two native Hebrew speakers were randomly assigned to the "Argentinian" guise (mean age = 28).

**Results:** As in the previous experiment, error rates were high in [h~x] trials. The same permissive criterion of inclusion was chosen, leaving 26 participants for the analysis.

The analysis of d-scores yielded a significant association in the predicted direction ( $t = 2.21, Df = 20, p = 0.039$ ). Block order had no effect ( $t = 1.33, Df = 12.22, p = 0.21$ ).<sup>12</sup> However, there was no significant difference between the d-scores found in 4a and those in 4b ( $t = 1.73, Df = 32.99, p = 0.093$ ). The distribution of d-scores by guise is presented in Figure 4.3.

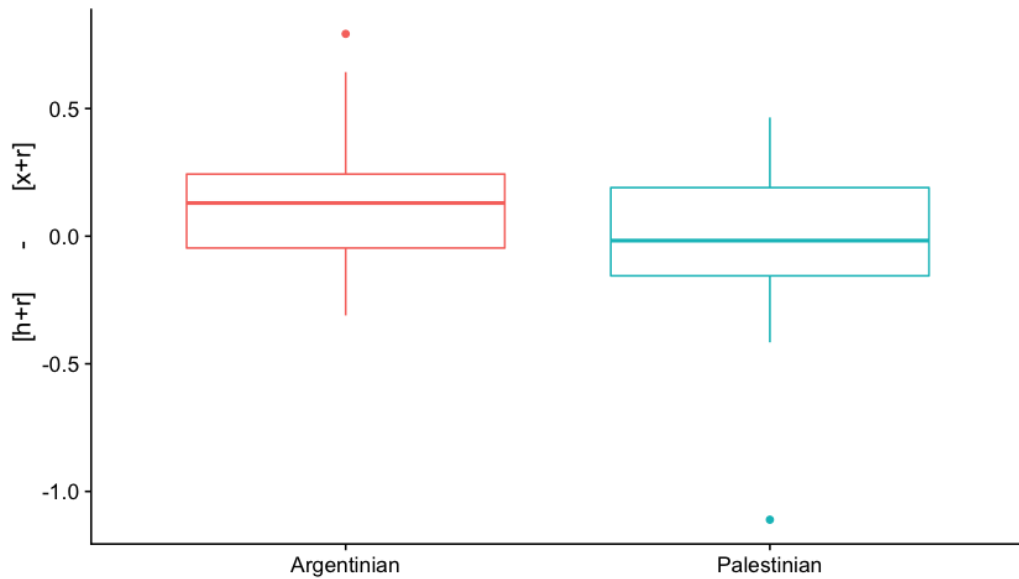
The overt judgements were significantly different from those obtained under the Palestinian guise: 27 participants marked [x] and [r] as related (the expected association under this guise), 7 supported an opposite view, associating [h] with [r], and 8 found no relationship between the variants. That is, the presentation of the guise significantly affected participants' views regarding the association between the variants in the predicted direction. A more detailed analysis of the open-ended questions is presented in the discussion.

#### 4.4.3 Success rates in categorization of target sounds in Experiments 5a+b

The d-score measure gave intermediate results: a significant effect in the expected direction, but only under the Argentinian guise. A more detailed post hoc analysis of the responses can potentially shed light on the opaque results.

---

<sup>12</sup> When participants with high error rates are included in the analysis, this effect disappears.



**Figure 4.3: Distributions of D-scores as a function of speakers' guise**

The stimuli used for [h]~[x] categorization were the same across guises, which allows for a direct comparison of the effect of guise on the categorization of these sounds. For this purpose, error rates and RTs were subjected to a generalized linear regression model and a linear mixed effects model, respectively. Only results from participants who had an overall success rate of above 75% were used in the analyses (the same dataset used for the d-score analysis). The fixed factors were Guise (Manuel/Waseem, sum-contrasted), Condition ([r]+[h] / [r]+[x], sum-contrasted), Order ([r]+[h] was presented first/second; sum-contrasted), and Stimulus ([r]/[h]/[x]; Helmert-contrasted, such that [r] trials were contrasted with both types of other trials and [h] and [x] trials were contrasted with each other). The Helmert contrast was chosen for the factor Stimulus Type for two reasons: (i) [r]-items were different recordings across guises, and (ii) the task can be viewed as two sub-tasks: association of [r] to either of the other variants; and for [h]/[x] items, the same association task in addition to phonetic categorization. That is why [r] can be viewed as a sort of baseline, that involves only one of the requirements in order to perform the task well.

Participant and Item were added as random intercepts. Mean error rates and RTs by guise, stimulus and condition are presented in Table 4.9. The full models are available in Appendix C; below I present only effects that were found to be significant.

The error rates model revealed a main effect of Stimulus Type, such that categorization of [x]~[h] was less accurate compared with categorization of [r] ( $\beta = -0.614$ ,  $SE = 0.08$ ,  $z = -7.67$ ,  $p < 0.0001$ ). No other main effects were found. Stimulus Type significantly interacted with Condition, such that categorization of [h] was more accurate in the [h]+[r] condition ( $\beta = 0.23$ ,  $SE = 0.078$ ,  $z = 2.99$ ,  $p < 0.003$ ).

Association	Stimulus	Error rate		Mean RT in ms (SD)	
		Palestinian	Argentinian	Palestinian	Argentinian
[r]+ [h]	[r]	1.8%	5.1%	1080 (353)	1171 (492)
	[h]	16%	9.3%	1307 (537)	1206 (529)
	[x]	14%	14%	1356 (538)	1393 (553)
[r]+ [x]	[r]	3.8%	2.6%	1109 (360)	1086 (386)
	[h]	16%	13%	1325 (559)	1213 (504)
	[x]	12%	11.5%	1356 (562)	1280 (513)

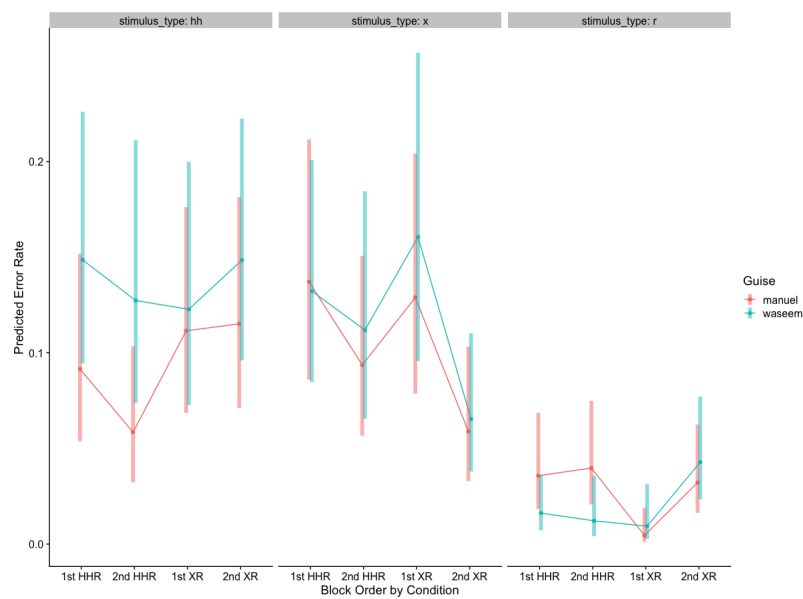
**Table 4.9: Mean error rates and reaction times by Guise, Stimulus and Condition of Experiment 5. Trials with RTs of above 2.5 standard deviations of the participant’s mean or below 150ms were discarded. Recall that the stimuli were identical across guises for [h] and [x], but not for [r].**

The Guise had no main effect on accuracy. However, there was a significant interaction between Guise and Stimulus Type, such that the categorization of [h] was more accurate in the Argentinian guise ( $\beta = 0.16$ ,  $SE = 0.08$ ,  $z = 2.06$ ,  $p = .04$ ). I analyze this result as stemming from a general effect of phonetic categorization bootstrapping under the Argentinian guise. That is, the difference in performance between the guises is due to increased attention to the [x] variant, which is used naturally by Spanish-Argentinian speakers (contrary to Arabic speakers). Participants under the Palestinian guise took longer to associate between the label used in the experiment (“soft Het”) and the target sound. Better convergence on the target category, prompted by the presentation of an alleged Argentinian speaker, reduced errors in categorization.

There was a marginal interaction of Guise with the Stimulus Type [r], such that the difference between [x]~[h] categorization and [r] categorization was smaller in the Argentinian guise ( $\beta = 0.18$ ,  $SE = 0.1$ ,  $z = 1.74$ ,  $p < 0.08$ ). This is the result of higher error rates with [r]-items under the Argentinian guise, an effect that can be attributed to the stimuli themselves (recall that [r]-items were recorded separately for each guise). There was also a marginal interaction of Guise with Condition, such that categorization in the [h]+[r] condition was overall more accurate under the Argentinian guise ( $\beta = 0.38$ ,  $SE = 0.225$ ,  $z = 1.678$ ,  $p = 0.09$ ). This effect is in the *opposite* direction of the expected association, since Argentinians typically produce [x] and [r]. The effect is composed of two triple interactions of Guise\*Condition\*Stimulus Type: in the [h]+[r] condition under the Argentinian guise, the difference between categorization of [r] (compared with categorization of [x]~[h]) was smaller compared with the overall difference between these stimulus types ( $\beta = 0.57$ ,  $SE = 0.21$ ,  $z = 2.79$ ,  $p = 0.005$ ). It reflects an overall *higher* error rate in [r] trials in the [h]+[r] condition under the Argentinian guise. This is the expected result in terms of association, since Argentinians typically produce [r] and [x]. Categorization of [h] (compared with [x]) was also marginally more accurate in the [h]+[r] condition under the Argentinian guise ( $\beta = 0.3$ ,  $SE = 0.155$ ,  $z = 1.955$ ,  $p = 0.051$ ). This finding may be due to the same phonetic categorization bootstrapping effect found in the Guise and Stimulus interaction, or alternatively, due to a higher error rates with [x]-items under this condition, which reflects more difficulty to associate [x] and [r] with different response keys under the Argentinian guise (i.e. the same predicted association

effect found with [r]).

Finally, while the order of blocks did not have a main effect, it interacted with Stimulus type: categorization of [h] compared with [x] improved between the first and second block ( $\beta = -0.54$ ,  $SE = 0.155$ ,  $z = -3.49$ ,  $p < 0.0005$ ), while the difference between categorization of [x]~[h] and categorization of [r] between the first and second block was reduced, i.e. [r] categorization was *worse* in the 2nd block ( $\beta = 0.65$ ,  $SE = 0.21$ ,  $z = 3.15$ ,  $p < 0.002$ ). I interpret this result as stemming from two parallel processes: improvement in [x]~[h] categorization on the one hand, and more errors in the association of [r] in the second block on the other, due to the change in response keys. There was also a triple interaction involving Stimulus Type, Block and Condition, in the opposite direction: when the [h]+[r] condition was presented first, categorization of [h] was improved more than when the [x]+[r] condition was presented first ( $\beta = -0.54$ ,  $SE = 0.155$ ,  $z = -3.49$ ,  $p < 0.0005$ ); the opposite was true with [r], which improved less between blocks when the [h]+[r] condition was presented first ( $\beta = 0.65$ ,  $SE = 0.21$ ,  $z = 3.15$ ,  $p < 0.0016$ ). Figure 4.4 is a visual summary of the accuracy model.



**Figure 4.4:** A visual description of the error rates model, by Stimulus Type (Left to right: [h], [x] or [r]), Condition ([h]+[r] vs. [x]+[r]), Order (1st vs. 2nd) and Guise (“Manuel” vs. “Waseem”). HH = [h]

The data are murky, but they can more easily be interpreted if we consider the two components of the task: categorization of [x]~[h] and association of [r]. The first task has been more challenging than anticipated, and the rate of success in this task depended on whether participants had a clear idea of what the target category is. When participants were not clear on the category, they tended to categorize [x]-items as [h]. When primed with the presentation of the Argentinian guise, participants could more easily converge on the target category [x], leading to fewer errors

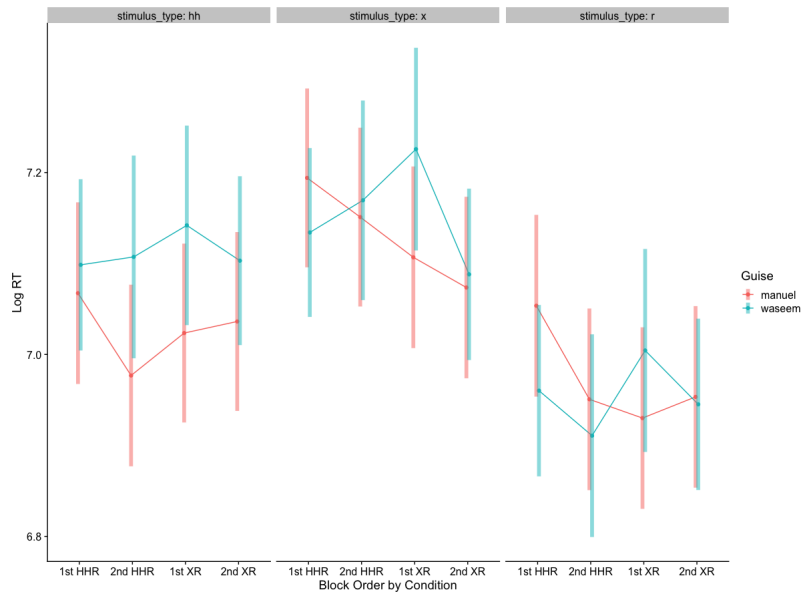
with [h]. The second task was simpler: associate the given category [r], that is phonetically easy to identify, with a response key. Two main issues should be taken into account here: first, the recorded stimuli were different between guises for this stimulus type, so the interpretation of the results should be within-guise ([r]-items cannot be compared across guises). Second, the [r] category is the only one that switches response key mid-task. Therefore, it is expected that errors would increase rather than decrease in the second block. The reaction times model is presented and analyzed below based on the same breakdown of the task.

The reaction times model revealed a main effect of Stimulus type, such that [x]-items were overall responded to more slowly compared with [h]-items ( $\beta = 0.04$ ,  $SE = 0.016$ ,  $Df = 12.06$ ,  $t = 2.34$ ,  $p < 0.04$ ). This may reflect the increased uncertainty of participants regarding this category's boundary. The Stimulus Type [r]-items were responded to faster compared with [x]~[h]-items ( $\beta = -0.05$ ,  $SE = 0.01$ ,  $Df = 12.23$ ,  $t = -5.23$ ,  $p < 0.001$ ). Order also had a main effect: participants were slower in the 1st block compared with the 2nd, indicating improvement in the task ( $\beta = 0.04$ ,  $SE = 0.006$ ,  $Df = 9585$ ,  $t = 6.2$ ,  $p < 0.0001$ ).

Condition had a marginal main effect: the [h]+[r] condition received overall slower responses ( $\beta = 0.012$ ,  $SE = 0.006$ ,  $Df = 9585$ ,  $t = 1.85$ ,  $p = 0.064$ ). Condition interacted with Stimulus type, such that [h]-items were responded to more slowly compared with [x] items under the Argentinian guise, but not under to Palestinian guise ( $\beta = 0.026$ ,  $SE = 0.0076$ ,  $Df = 9585$ ,  $t = 3.45$ ,  $p < 0.0001$ ). Condition also interacted with Guise, such that responses in the [h]+[r] condition were slower in the Argentinian guise ( $\beta = 0.066$ ,  $SE = 0.013$ ,  $Df = 9585$ ,  $t = 5.203$ ,  $p < .0001$ ). This interaction is expected under the main hypothesis of the study, since Argentinians typically produce [r] and [x].

There was no main effect of Guise on RT, but in addition to the interaction with Condition, it was found that the Guise also significantly interacted with Stimulus type (as found previously in the accuracy model). Specifically, [h]-items got significantly longer RTs than [x] in the Argentinian guise, but comparable RTs in the Palestinian guise ( $\beta = 0.032$ ,  $SE = 0.008$ ,  $Df = 9585$ ,  $t = 4.202$ ,  $p < 0.0001$ ). This mirrors the effect of more accurate categorization of [x]~[h] under the Argentinian guise, lending further support to the proposal that the Argentinian guise helped participants focus on the relevant acoustic dimension. There was also an interaction of Guise and the contrast between [r] and [x]~[h] categorization, such that [r]-items were responded to more slowly under the Argentinian guise ( $\beta = 0.023$ ,  $SE = 0.0046$ ,  $Df = 9585$ ,  $t = 5.148$ ,  $p < 0.0001$ ). This result is probably mostly due to acoustic properties that distinguish between the [r]-items in the different guises, which relied on separate recordings.

There were also triple interactions, involving all fixed factors: Stimulus Type, Guise, Condition and Order. Guise interacted with Stimulus type and Condition following a similar pattern to that found for the accuracy rates: Categorization of [h] (compared with [x]) was faster in the [h]+[r] condition in the Argentinian guise ( $\beta = 0.038$ ,  $SE = 0.015$ ,  $Df = 9585$ ,  $t = 2.492$ ,  $p = 0.013$ ). I interpret this finding as evidence that it was more difficult to respond to [x] when it was not associated with the same response key as [r] under the Argentinian guise (Spanish-Argentinian speakers produce [r] and [x]). RTs for [r]-items were marginally slower in the [h]+[r] condition



**Figure 4.5: A visual description of the RTs model, by Stimulus Type (Left to right: [r], [h] or [x]), Condition ([h]+[r] vs. [x]+[r]), Order (1st vs. 2nd) and Guise ("Manuel" vs. "Waseem"). HH = [h]**

under the Argentinian guise ( $\beta = 0.017$ ,  $SE = 0.009$ ,  $Df = 9585$ ,  $t = 1.799$ ,  $p = 0.072$ ) – the expected result in terms of the association.

Stimulus Type, Block and Condition also interacted as in the accuracy model: when the [h]+[r] condition was presented first, categorization of [h] was improved more than when the [x]+[r] condition was presented first ( $\beta = -0.055$ ,  $SE = 0.015$ ,  $Df = 9585$ ,  $t = -3.608$ ,  $p < 0.0003$ ); that is, RTs for [h]-items became shorter between the 1st and 2nd blocks when the [h]+[r] condition was presented first. The opposite was true with [r], which improved less between blocks when the [h]+[r] condition was presented first ( $\beta = 0.028$ ,  $SE = 0.009$ ,  $Df = 9585$ ,  $t = 3.058$ ,  $p < 0.002$ ). Figure 4.5 is a visual summary of the RT model.

Taken together, the results suggest that it has overall been more difficult to distinguish between [h] and [x] in the Palestinian guise, despite the fact that the stimuli for [x]~[h] categorization in both guises were identical. As mentioned through the presentation of the results, I take effects with [r]-items that are not within-guise (i.e. that involve a comparison between different stimuli) to be a result of the recordings themselves. However, the fact that RTs for [r]-items were shorter in the Palestinian guise with the [h]+[r] association and in the Argentinian guise in the [x]+[r] association indicates a trend in the predicted direction of association. The relative difficulty of the phonetic categorization task obscured the results for association in [x]~[h] categorization.

#### 4.4.4 Discussion

Experiment 5 revealed an implicit association between the target variants in the predicted direction under the Argentinian guise, and no association under the Palestinian guise. No significant difference in the strength of associations between the guises was found.

Post-hoc analyses were conducted with the goal of providing an explanation of this intermediate result. Three significant differences in performance based on the Guise were obtained: the explicit association reported by listeners, the performance in [h]~[x] categorization, and the increased error rate as well as longer RTs in Guise-incongruous [r] trials.

The explicit association between variants differed based on guise: [h] and [r] were associated by more participants in the Palestinian guise, and [x] and [r] were associated by more participants under the Argentinian guise, as expected. This result suggests that the guise was salient enough in context and familiar enough to participants.

Error rates in categorization were high under both guises, suggesting that it was hard for participants to distinguish between [h] and [x]. The difficulty was augmented under the Palestinian guise. This result is not trivial, since the stimuli were identical. In Section 4.4.3 above, I proposed that the difference in performance between the guises is due to increased attention to the [x] variant, affected by the presentation of the Argentinian guise. The category of [x] is used by (and in media portrayals of) at least two different personae: L1 Spanish and L1 Russian speakers. Thus, it is not exclusively associated with Spanish speakers. Unlike the [h] variant, that is represented in the speech of native as well as non-native speakers and has a name in Hebrew, the [x] variant seems not to be widely stereotyped for Hebrew speakers. When listeners were presented with a persona that is known to use [x], they could more easily converge on the phonetic distinction between the acoustically similar segments [x] and [h]: since the boundaries of the category [x] became clearer under the "Argentinian" guise, listeners erred less in [h] trials.

It is important to note that other cues within [h]/[x] words could also have contributed to the better recognition of the target sounds, despite not being stereotyped. This is because, although there were no known indicators of an Arabic/Spanish accent in the recordings, the speaker may have "carried" prosodic features that are congruent with each target segment, thus contributing more to the "Argentinian-ness" or "Arab-ness" of the item overall. Further, as is customary in the IAT, each category was represented by very few items (5 in the current case), so listeners can rely on idiosyncratic features of the recordings for categorization. It is likely that listeners who were primed by the Argentinian guise could also rely more on other features of the recording, responding to "Argentinian-ness" and "Arab-ness" instead of the target sounds. This account is congruent with the results of /-ing/-variant recognition in English: as Vaughn and Kendall (2018) found, listeners could perform well above chance in guessing whether the used variant of 'ing' – that was masked during presentation – was alveolar or velar, based on subtle phonetic cues within the rest of the sentence (see Section 4.1). This proposal does not necessarily contradict the previous, based on category boundaries alone; both the focus on category boundaries and

the overall phonetic composition of items could have contributed to the current results, and both explanations rely on speaker identification as the cause.

The proposal that the guise affected processing is further strengthened by the answers of participants to the open questions about their overt judgements of each variant. When asked “What social groups do you know who uses the variant?”, participants were more likely to associate [r] with Spanish in the Argentinian guise, and more likely to associate it with Arabic in the Palestinian guise. Similarly, [x] was more likely to be associated with Spanish in the Argentinian guise, although there was much more ambiguity associated with this variant than with the others: under both guises, almost half of the participants mistook [x] to be the unmarked variant [χ], responding to the same post-task question with the answer “most young native Hebrew speakers use this variant”.

Table 4.10 summarizes participants’ answers to the question, by guise. All participants were included in this analysis, including participants whose data were removed due to high error rates. Some participants did not answer the questions in each guise, leaving a total of 42 participants for the Argentinian guise and 40 in the Palestinian guise. In their answers, participants could mention as many social groups as they wanted, hence the percentage don’t sum up into 100% (the percentage represents the total of participants that mentioned a particular group as using the given variant).

Argentinian guise (“Manuel”)			
	[h̥]	[r]	[x]
Arabic	97%	26%	4.8%
Spanish	2.4%	95%	38%
Native Hebrew	0	0	45.2%
Other	0	42.8%	23.8%
Palestinian guise (“Waseem”)			
	[h̥]	[r]	[x]
Arabic	97.5%	55%	12.5%
Spanish	0	40%	12.5%
Native Hebrew	0	0	47.5%
Other	0	45%	25%

**Table 4.10: Overt judgments of participants in Experiment 5: the rate of association between a given social group and the target variants. Other = mostly Russian and other European languages.**

#### 4.5 General discussion

The current chapter explored two case studies of co-variation in Modern Hebrew: [h̥] with [ej]~[e] (Experiment 4) and [r] with [h̥]~[x] (Experiment 5). Experiment 4 revealed a medium to strong association between [h̥] and [ej], with no mediating effect of the guise. The results of Experiment 5 were intermediate: on the one hand, there was no single association pattern, irrespective of guise,



as in Experiment 4. On the other, there was a significant association between [r] and [x] in the Argentinian guise, but not the corresponding association between [r] and [h] in the Palestinian guise.

The explicit judgments were congruent with the intended guise: more participants associated between [h] and [r] in the Palestinian guise than in the Argentinian guise, and the opposite was true with [x] and [r]. In light of the inconclusive results with respect to the main question (namely, whether guise can affect automatic processing of variants), the discussion below is dedicated to the following related questions:

1. What type of representations are targeted by the IAT?
2. Why weren't the explicit judgments of speakers correlated with their results in the IAT?

The result that the social context, i.e. the guise presented to participants, did not significantly mediate the association between linguistic markers in the IAT, is surprising given previous results from the same paradigm. For example, Dasgupta and Greenwald (2001) demonstrated that the IATs of White Americans showed less bias against Black people after being exposed to positively-viewed Blacks. Blair et al. (2001) showed a similar effect with Women and weakness-related attributes: after imagining strong women, participants' d-score in the IAT dropped. These two examples used specific exemplars from a social group to counter the stereotypes against it. Another type of successful contextual intervention targeted the meta-categories themselves: Foroni and Mayr (2005) used a story that changed the very attributes of two objects, insects and flowers. In their story, presented to participants before the IAT task, humans in a post-apocalyptic world survive on eating insects and avoid flowers, that are highly toxic. The story caused a reversal of the IAT effect in 28% of the cases.

A similar modulation of the IAT results was obtained with a bi-modal version of the IAT, in which phonological variants were presented in the auditory modality and attributes were presented orthographically. Weirich et al. (2020) presented listeners with the same stimuli that is associated with two distinct social groups: L1 French speakers and "Kiezdeutsch" speakers, a variety of German spoken mainly by second-generation immigrants in the Berlin area. The marked variant, identical in both social groups, was presented as "L1 French" in one condition and as "Kiezdeutsch" in the other. They found that the category labels themselves mediated the results for some groups of participants, such that the d-scores of old Germans, reflecting aversion, were higher when the variant was perceived as a feature of second generation immigrant than when it was perceived as a feature of a German-learning L1 French speaker (young participants were not affected by the category labels).

In sum, IAT scores are context sensitive. In addition, it has previously been demonstrated in other paradigms that the identity of the speaker affects the processing and categorization of speech sounds (Strand 1999; Phrao et al. 2014; D'Onofrio 2018a). Why, then, weren't the implicit

associations between variants altered based on the perceived persona of the speaker? I hypothesize that the key to understanding this apparent discrepancy is the kind of mental structures targeted by the IAT, and the difference between visual and auditory stimuli within the paradigm.

First, in all of the above designs, the relationship is between an object (social groups, or linguistic markers of social group in the last example) and attitudes/attributes (good and bad). In the current study, auditory variants were associated with other auditory variants. The manipulation of attitudes on the basis of a short story involves the salient psychological constructs of “good” and “bad”.<sup>13</sup> In contrast with these high-level constructs, linguistic variants often mark more than one social group or speech style, and thus have more complex relationships with other variants. This ambiguity was manifested in the disagreement with regard to the association of the variants [r] and [x] with one social category (see Table 4.10). While there was a general agreement among participants that the [h] variant characterized L1 Arabic and Mizrahi speakers (more than 97% under both guises), [x] was associated with native Russian speakers by about 25% of the participants in both guises, while almost 50% of the participants did not distinguish between this variant and the variant used by the majority dialect of Hebrew, [χ]. [r] was associated with L1 Spanish speakers by 95% of the participants in the Argentinian guise, but only by 40% of the participants in the Palestinian guise; it was associated with L1 Arabic speakers by 55% of the speakers in the Palestinian guise, but only by 26% of the participants in the Argentinian guise.

A related point is that with auditory presentation, the IAT becomes more opaque: listeners can tie variables into ad-hoc meta-categories, such as “native/non-native”, “marked/unmarked”, “orthographically transparent/opaque”, “reduced/full”. Since we cannot know in which direction participants were pulled, the results become less transparent as well. As I proposed in the discussion of Experiment 4, in some cases, participants might prefer to ignore the persona and rely on a different organizing principle (in the case of Experiment 4, socially marked vs. unmarked variants). The explicit judgments, on the other hand, allowed participants direct access to meta-linguistic control. The explicit judgments were collected after the IAT, without a time limitation. This gave participants substantial time to reflect on the relationship between the variants in light of the persona they were introduced with. In this context, the question “which of the variants presented in the experiment are related in your opinion?” is a question about *given categories*. It does not require participants to categorize items themselves, a harder and more ambiguous task, as discussed above. When it comes down to meta-linguistic control alone, participants were affected by the guise they were presented with.

Second, auditory IATs are more difficult for participants. As they unfold over time and do not remain on the screen, auditory stimuli require more attention than reading, as evident in the longer RTs and higher error rates for auditory stimuli compared with visual ones within experiments in which both types of stimulus were presented (Campbell-Kibler 2012; Weirich et al. 2020). This difficulty can obscure the results: participants make more errors and respond more slowly, perhaps

---

<sup>13</sup> It is interesting to consider the reduction of bias in old Germans, observed by Weirich et al. (2020), as a change of the organizing principle of the meta-category: prestigious L2 speech vs. non-prestigious native speech.

obscuring the effect of the mismatch condition, even assuming that participants indeed associate the variant with the relevant social group. This difficulty can increase dramatically if, as in the case of Experiment 5, the acoustic difference between variants is delicate. That is, the IAT relies on good categorization in order to test associations; the assumption that phonetic categorization is at a good enough level should be affirmed before the task can be used.

#### **4.6 Conclusions**

The IAT explores implicit relations between abstract objects. In the context of linguistic variation, it has previously been demonstrated that social markers that belong to the same dialect show a stronger association with each other compared with markers that do not belong to the dialect.

Experiment 4 revealed an association between two marked variants, [h̃] and [ej]. Experiment 5 revealed an association between [r] and [x] in the Argentinian guise. However, there was no significant difference between d-scores based on guise in either experiment. In Experiment 4, the association found between the two variants seemed general and was not affected by Guise. In the discussion, I proposed that this result is due to the socially marked status of both [h̃] and [ej], that overshadowed the persona manipulation.

The intermediate results of Experiment 5 were attributed to the difficulty to make the [h̃]~[x], due to (i) acoustic similarity; (ii) unfamiliar label for [x]. The Palestinian guise was more acutely affected by (ii), since the alleged "Argentinian" speaker made available the category [x], thus focusing participants attention on its particular acoustic features, allowing them to distinguish it from [h̃].

Despite the lack of a reliable difference in d-scores, Experiment 5 showed two effects of online speaker-oriented processing, in addition to the offline effect of explicit judgments. First, in the Argentinian guise, which had overall higher accuracy rates, there was a significant association in the expected direction. Second, categorization was sensitive to the speaker: when the acoustic difference between categories was subtle ([h̃]~[x]), exposure to the guise that is known to produce the less-known category bootstrapped categorization.

## V

# Conclusions

This thesis explores the perception of ethnically-conditioned phonological variation: its social meaning and its implications on auditory speech processing. Part I enhances our understanding of the social perception of phonological variation in Modern Hebrew, and Part II utilizes socially meaningful phonological variation in order to study cognitive mechanisms of speech processing, with an emphasis on how perceived social identity may interfere with word recognition. In the following paragraphs, I highlight the contribution of the current research to the study of variation in Modern Hebrew, longstanding questions about phonological representations and the field of online speech processing.

### ***5.1 Implications for variation studies on Modern Hebrew***

The current work provides a new framework for understanding the retention of Mizrahi features of speech in Modern Hebrew. While most previous studies emphasize the process of convergence towards an Ashkenazi-led norm, I propose that there is no general process of convergence that applies evenly throughout the non-Ashkenazi population, but rather a series of individual-level tendencies to conform with a non-stereotyped norm upon moving to an Ashkenazi-majority environment. Due to the controlled dispersion of the population, whereby communities were settled in ethnically-homogeneous neighborhoods, development towns and villages, first generation Mizrahi speakers in the periphery retained and developed features that originate from their L1, primarily Arabic. These features, which include lexical, syntactic and phonological characteristics, are still prevalent in the same areas of the state of Israel, forming a variety of Hebrew sometimes termed Mizrahi (Matras and Schiff 2005), Peripheral (Henshke 2013b), or in some cases “Working class/low SES” Hebrew.

In Chapter 1, I argued that the variationist literature on Modern Hebrew suffers from a persisting bias, which caused many authors to overlook variation, report its decline, and later attribute most forms of variation to socio-economic status or individual deficiency. This bias is at least in part rooted in the iconization (in the sense of Irvine and Gal 2000) of the language as a symbol of the resurrecting nation, which is a linguistic reflection of the Melting Pot doctrine enforced during the first decades of the state. Documented non-professional debates on variation from the periods in which the first variationist works were published, as well as the geo-historical circumstances

and, most importantly, the status of variation today, all suggest that variation is not undergoing an across-the-board decline. Instead, I argue, some forms of variation gradually disappear due to intense stereotypization, while others continue to thrive in communities in which the Mizrahi sociolect is default. More generally, segregation between communities and homogeneity within them is associated with increased linguistic variation, not convergence. In this context, it also seems that some young speakers of an Ashkenazi-majority background (in specific geographic areas) diverge in their productions in new recognizable ways.

First empirical evidence supporting this claim are found in Chapter 2: as Experiment 1 demonstrated, young native speakers of Modern Hebrew were perceived as significantly different from each other in terms of ethnicity in a consistent way, based on recordings of  $\sim 5$  seconds each. The results of Experiment 1 showed that sounding Ashkenazi is not always a default. Specifically, some speakers were not consistently rated as sounding Mizrahi or Ashkenazi, while others were, strengthening the view proposed in Chapter 1 that neither of the varieties should be considered an unmarked default.

The acoustic analysis revealed a significant correlation between rhythm and ethnicity perception. Since the acoustic analysis only tested specific features of the recordings, it is not possible to determine that rhythm was the factor that caused the divergence in evaluations; it might be the case that other features, that rhythm happens to be correlated with, are the cause. While much work is still needed in order to characterize which cues are relevant for the evaluation of speakers' ethnicity, these preliminary findings provide a new starting point for future investigations. Certainly, these findings contribute the insight that phonological variation in Hebrew is socially evaluated in a similar way by native listeners.

In sum, it was claimed that the common view – that variation in Modern Hebrew is on continuous decline and will ultimately be eliminated – should be cleared off in favor of a detailed examination of existing variation and the perspectives of Hebrew speakers. There is reason to assume that there is more than one default variety of Modern Hebrew, and so, variationists should document linguistic features of speakers' communities that perceive themselves as distinct, based on cultural, ideological, geographic and ethnic factors. More than anyone, researchers in cognitive development and education studies should be aware of multiple production norms when assessing children's syntactic and morphological development.

## **5.2 Implications for phonological theory**

Part II explores related questions about the interaction between social and phonological perception in online speech processing, using quantitative experimental methods. It utilizes known social markers, and stereotyped personae who produce these markers, in order to assess the phonological representations of speakers of various linguistic backgrounds on the one hand, and the role of speaker-oriented processing on the other.

Experiments 2-3 (Chapter 3) studied the processing and representations of the marked pharyngeal [ħ] by speakers from Merged vs. Pharyngealizing backgrounds. This marker has merged with [χ] in the speech of most young native speakers of the language, including speakers who are second generation to Pharyngealizing parents (“Pharyngeal listeners”). Pharyngealizers and Pharyngeal listeners rejected switched  $\chi \rightarrow \text{ħ}$  words at a much higher rate than Merged listeners (who are second generation listeners to the Merged variety). Merged listeners could still distinguish between the categories at a near-native level (Experiment 2), indicating that the lexical, and not phonetic encoding of the categories is the main source of the divergent pattern between the groups. This result is another illustration that some divergences in phonological encoding stem from an assimilation process at the level of the lexicon: as studies on English, Dutch and French L2 learners have shown previously, there is an inherent a-symmetry in encoding between the L2 category that is phonetically more similar to some L1 category, and the “new” sound, which is perceived as a poor fit (or ‘bad exemplar’) of the same category (Weber and Cutler 2004; Darcy et al. 2013; Melnik and Peperkamp 2019).

The fact that Merged speakers do not distinguish between [ħ] and [χ] in their lexicon also bears on a longstanding debate in generative phonology. Since the *Sound Patterns of English* (SPE; Chomsky and Halle 1968), the book that presented what later became considered the classical linear rule theory of generative phonology, the field is ambivalent regarding the level of abstractness that representations in phonology may reach. The SPE was highly abstract. Underlying representations (URs; the forms stored in the lexicon) were sometimes very different from the phonetic representations (PRs), to a degree that *some URs never appeared on the surface*. Kiparsky (1968) argued that the level of abstractness reached in the SPE would raise serious problems for the learners of a language, and proposed a condition against some relationships between UR and PR:

**The Alternation Condition:** An underlying form has to appear on the surface in at least one context. There are no elements in the underlying representation which never surface.

The Alternation Condition rules out cases dubbed “absolute neutralization”, i.e. cases in which an underlying form *never* surfaces. Yet, for Merged listeners, this is the case with the marked pharyngeal [ħ]: the sound never surfaces in the PRs they are exposed to (as speakers in their close environment merge the categories), but there are still systematic phonological differences between “non-alternating [χ]” and “alternating [χ]~[k]”, which stems (under an SPE-style approach) from their underlying representations as [ħ] and [χ], respectively. In an alternative approach, these systematic differences are ignored at the level of grammar and memorized to the lexicon.

As mentioned in Chapter 3, [ħ] and [χ] trigger different phonological alternations. For example, underlying /ħ/ causes the lowering of [e] to [a]. Compare the Merged-variety words *oreχet* ‘editor’ and *oraxat* ‘guest’, within the same template. The latter [χ] was historically [ħ]. Another example: [χ] may alternate with [k], while /ħ/ never surfaces as [k] (kitev - jeχatev ‘send a copy of a correspondence’ past vs. future, χibek - jeχabek ‘hug’ past vs. future). Chomsky and Halle (1968) assumed that differences like these in phonological behavior can establish the creation of

two separate categories in the minds of native Hebrew speakers, despite the identical phonetic form of the two categories. Kiparsky (1968) claimed that it cannot be the case – in order to establish that categories differ from each other, there must be *phonetic* (surface) evidence. Experiment 3 supports this conclusion: merged listeners, who were not exposed to overt phonetic evidence that the categories are distinct, seem not to represent them as distinct in the lexicon. Further, even evidence from the orthographic system, in which each of these categories is represented by a different letter, did not suffice.

It is worth mentioning that at the same time, the “Pharyngeal listeners” group embody a situation of “Absolute Neutralization”: their merged production does not match their alleged underlying representation. Such a linguistic situation is probably destined to be short-lived, and occur only during times of linguistic changes in progress; in the current case, this intermediate generation includes pharyngeal listeners who are merged speakers. After all, the children of Pharyngeal listeners will receive the same input as those of Merged listeners, because both groups produce the Merged variant.

An alternative interpretation can redeem the SPE-style abstract phonology. The real word stimuli in Experiment 3 included only nouns, that often have no clear paradigm, and therefore have less related words in which [k] stands for [χ]. In verbs, the paradigm is much more clear and used more often (e.g. all the past vs. future forms with /k/ in the initial consonant position involve the stop-fricative alternation). The verbal paradigm thus presents an opportunity for a follow-up study to tease apart these accounts; it is possible that in the verbal context, Merged listeners would be more reluctant to accept switched-category words, relying on the related words in which [χ] alternates with [k] – a result that would grant support to a more abstract representation of such items.

### **5.3 Implications for online speech processing in the indexical route**

Apart from the implications regarding the mental representations of speakers from different linguistic backgrounds, Experiment 3 also revealed an effect of speaker-oriented processing: listeners from a pharyngealizing background were more likely to reject (Pharyngealizers) or respond significantly more slowly (Pharyngealizers and Pharyngeal listeners) to switched  $h \rightarrow \chi$  words when produced in the Pharyngealizing speaker’s voice. This is the condition dubbed “Inconsistent”, which is on the one hand a common production of the given item (indeed, this is its production in the Merged variety), and on the other, inconsistent with the identity of the speaker as presented so far. This finding illustrates that deep acquaintance with a vernacular can induce subtle predictions regarding production; even when the produced form was not illicit at the global (lexical) level, but only inconsistent with the identity of the speaker, it was rejected or responded to more slowly by highly experienced listeners. For the model presented in Cai et al. (2017), in which linguistic inputs are processed simultaneously in the lexical and indexical (speaker-oriented) route, this finding provides another piece of evidence in favor of the hypothesis that speaker-related information affects processing in real time.

Chapter 4 targeted more specifically this type of online speaker-oriented processing effect, using the Implicit Association Task (Greenwald et al. 1998). The goal of the experiment was to test whether the association between two marked variants can be modulated by the perceived identity of the speaker. This goal was not achieved using the d-score algorithm, used to assess the magnitude of association within the paradigm: In Experiment 4, both personae yielded an association in the same direction and of a comparable magnitude; in Experiment 5, a significant association in the predicted direction was found only under one of the guises, and there was no significant difference between them. A great deal of the general discussion of this chapter examined these results in the context of previous IATs. The upshot is a methodological recommendation to avoid studying linguistic variants that are not both highly salient and socially marked with the IAT, among other things because the assessment of the magnitude of association is based on the assumption that the baseline of categorization is good.

Despite the inability to detect a significant difference in d-scores based on the guise, post hoc analyses of accuracy and RTs revealed an effect of the expected reversal of associations in Experiment 5. Two findings are of particular interest to the dual-route model of speech processing presented in Cai et al. (2017): (i) Categorization of [h] was more accurate under the Argentinian guise. (ii) RTs were significantly longer in the [h]+[r] condition under the Argentinian guise (i.e. the expected direction of association; the variants used in Spanish-accented speech are [x] and [r], while Arabic-accented speech includes [h] and [r]). This is congruent with the d-score under the Argentinian guise, which revealed an association between [x] and [r].

The proposed interpretation of finding (i) was that [x] was not acoustically salient enough for listeners, and given that there is no familiar and unified label that denotes this category in Hebrew (contrary to the other marked variants used in the task), many participants could not converge on the intended category. Under the Argentinian guise, the category became more salient; participants could couple their meta-linguistic knowledge (“Argentinians tend to produce recognizable [x] sounds, that differ in such-and-such way from the received pronunciation in Hebrew”) with the new label name (“Soft Het”), thus converging on the relevant acoustic dimension that allowed them to distinguish between [x] and [h], leading to higher accuracy in [h] trials. This view was supported by the overt judgments of participants and the post-task questionnaire: under the Argentinian guise, participants were more likely to associate [x] with an Argentinian accent.

This finding is not trivial. It illustrates that meta-linguistic knowledge can help focus listeners’ attention on a relevant acoustic dimension. It also provides empirical support for an intuition shared by listeners in a multi-lingual environment: in situations in which it is not clear in which language one is being addressed, speech decoding becomes more laborious. Finding (ii) suggests that the Guise affected the ease of association between variants. It has previously been demonstrated that the gender of a speaker and their persona can affect category boundaries (Strand 1999; D’Onofrio 2018b). However, these results could have been due to late meta-linguistic ‘calibration’. That is, a listener can think something along the lines of “this sounds to me more like an /a/, but the speaker is supposed to be a valley girl, so it’s probably an /æ/” – and respond accordingly in the categorization task. There seems to be no meta-linguistic reasoning that would explain longer



RTs or higher error rates for the same stimuli under a particular guise. Thus, the results are likely to reflect an online effect of calibration, that can only be attributed to the perceived identity of the speaker.

In sum, Part II supports the model proposed in Cai et al. (2017), suggesting that social information based on perceived speaker identity contributes to online speech processing: Chapter 3 demonstrates that deep acquaintance with a vernacular can induce subtle predictions regarding production, such as whether the speaker would produce a merged category; Chapter 4 adds that the presentation of a speaker who uses a specific variant can focus listeners' attention on the relevant acoustic dimension for categorization, as well as affect the predictions regarding the distribution of other variants in that persona's speech.

## References

- Adank, Patti, Bronwen G Evans, Jane Stuart-Smith, and Sophie K Scott. 2009. Comprehension of familiar and unfamiliar native accents under adverse listening conditions. *Journal of Experimental Psychology: Human Perception and Performance* 35:520.
- Azriyahu, Maoz. 2000. Hebrew and hebraicization in the formation of cultural identity [ivrit ve-ivrut: hebetim šel jecirat zehut tarbutit]. *Jewish Studies* 77–88.
- Babel, Molly. 2012. Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics* 40:177–189.
- Babel, Molly, Grant McGuire, Sophia Walters, and Alice Nicholls. 2014. Novelty and social preference in phonetic accommodation. *Laboratory Phonology* 5:123–150.
- Barkat, Melissa, John Ohala, and Francois Pellegrino. 1999. Prosody as a distinctive feature for the discrimination of arabic dialects. In *Sixth European Conference on Speech Communication and Technology*.
- Bates, Douglas, Deepayan Sarkar, Maintainer Douglas Bates, and L Matrix. 2007. The lme4 package. *R package version* 2:74.
- Ben David, Abba. 1974. *Pronunciation guide to radio and television (hebrew: madrix lashon le-radio u-le-televizia)*. Broadcasting Authority Training Center.
- Ben Tolila, Yaakov. 1984. The pronunciation of hebrew: Social phonology. *Jerusalem: Hebrew University [in Hebrew]* .
- Ben Tolila, Yaakov. 2002. *Speaking hebrew: Studies in the spoken language and in linguistic variation in israel [in hebrew]*, chapter Linguistic variation across generations in Israel, 131–140. (Te'uda 18).
- Benger Alaluf, Ya'ara, and Noa Bassel. 2021. The tel aviv syndrome: when housing shortage meets national-religious segregation (syndrom tel aviv: kshe-metsukat dijur pogeshet badlanut le'umit datit [hebrew]). URL <https://www.haokets.org/2021/05/06/%D7%A1%D7%99%D7%A0%D7%93%D7%A8%D7%95%D7%9D-%D7%AA%D7%9C-%D7%90%D7%91%D7%99%D7%91-%D7%9B%D7%A9%D7%9E%D7%A6%D7%95%D7%A7%D7%AA-%D7%93%D7%99%D7%95%D7%A8-%D7%A4%D7%95%D7%92%D7%A9%D7%AA-%D7%91%D7%93%D7%9C/>.

- Blair, Irene V, Jennifer E Ma, and Alison P Lenton. 2001. Imagining stereotypes away: the moderation of implicit stereotypes through mental imagery. *Journal of personality and social psychology* 81:828.
- Blanc, Haim. 1968. *The israeli koine as an emergent national standard*. John Wiley & Sons, Incorporated.
- Boersma, Paul. 2006. Praat: doing phonetics by computer. <http://www.praat.org/> .
- Bolozky, Shmuel. 2002. Šonut fonologit u-morfologit ba-‘ivrit ha-meduberet [‘phonological and morphological variation in spoken hebrew’]. *Izre’el*, ed 239–278.
- Cai, Zhenguang G, Rebecca A Gilbert, Matthew H Davis, M Gareth Gaskell, Lauren Farrar, Sarah Adler, and Jennifer M Rodd. 2017. Accent modulates access to word meaning: Evidence for a speaker-model account of spoken word recognition. *Cognitive Psychology* 98:73–101.
- Campbell-Kibler, Kathryn. 2007. Accent,(ing), and the social logic of listener perceptions. *American speech* 82:32–64.
- Campbell-Kibler, Kathryn. 2012. The implicit association test and sociolinguistic meaning. *Lingua: International Review of General Linguistics* 122:753–763.
- Chéreau, Céline, M Gareth Gaskell, and Nicolas Dumay. 2007. Reading spoken words: Orthographic effects in auditory priming. *Cognition* 102:341–360.
- Chomsky, Noam, and Morris Halle. 1968. *The sound patterns of english*. Harper & Row New York.
- Clopper, Cynthia G, and Ann R Bradlow. 2008. Perception of dialect variation in noise: Intelligibility and classification. *Language and speech* 51:175–198.
- Cohen, Evan-Gary, Lior Laks, and Carmen Savu. 2019. The phonetics of modern hebrew rhotics. *Brill’s Journal of Afroasiatic Languages and Linguistics* 11:28–48.
- Cohen, Yinon. 1998. Socioeconomic gaps between mizrahis and and ashkenazis, 1975-1995. *Sociologia Israelit* 117–173.
- Conrey, Brianna, Geoffrey F Potts, and Nancy A Niedzielski. 2005. Effects of dialect on merger perception: Erp and behavioral correlates. *Brain and Language* 95:435–449.

- Dahan, Momi, Eyal Dvir, Natalia Mironychev, and Samuel Shye. 2003. Have the gaps in education narrowed? on factors determining eligibility for the israeli matriculation certificate. *On Factors Determining Eligibility for the Israeli Matriculation Certificate* .
- Dahan, Yossi, and Yossi Yona. 2013. The israeli education system: equality and opportunities - from building a nation to neo-liberalism. In *Practices of differentiation in the education field in israel [hebrew]*.
- Darcy, Isabelle, Danielle Daidone, and Chisato Kojima. 2013. Asymmetric lexical access and fuzzy lexical representations in second language learners. *The Mental Lexicon* 8:372–420.
- Dasgupta, Nilanjana, and Anthony G Greenwald. 2001. On the malleability of automatic attitudes: combating automatic prejudice with images of admired and disliked individuals. *Journal of personality and social psychology* 81:800.
- Dauer, Rebecca M. 1983. Stress-timing and syllable-timing reanalyzed. *Journal of phonetics* 11:51–62.
- Davis, Lawrence M. 1984. The pharyngels in hebrew: Linguistic change in apparent time. *Folia Linguistica Historica* 18:25–32.
- De Mareüil, Philippe Boula, and Bianca Vieru-Dimulescu. 2006. The contribution of prosody to the perception of foreign accent. *Phonetica* 63:247–267.
- Devens, Monica S. 1981. Misconceptions about accent and national origin among native israeli hebrew speakers: A preliminary report .
- D’Onofrio, Annette. 2018a. Controlled and automatic perceptions of a sociolinguistic marker. *Language Variation and Change* 30:261–285.
- D’Onofrio, Annette. 2018b. Personae and phonetic detail in sociolinguistic signs. *Language in Society* 47:513–539.
- Dufour, Sophie, Noël Nguyen, and Ulrich Hans Frauenfelder. 2007. The perception of phonemic contrasts in a non-native dialect. *The Journal of the Acoustical Society of America* 121:EL131–EL136.
- Eckert, Penelope. 2000. *Language variation as social practice: The linguistic construction of identity in belten high*. Wiley-Blackwell.

- Eckert, Penelope. 2008a. Variation and the indexical field 1. *Journal of sociolinguistics* 12:453–476.
- Eckert, Penelope. 2008b. Where do ethnolects stop? *International journal of bilingualism* 12:25–42.
- Even-Zohar, Itamar. 1979. The growth and consolidation of local native culture hebrew in the land of israel, 1882-1948 (ha-tsmiha ve-ha-hitgabshut shel tarbut ivrit mekomit ve-yelidit be-eretz israel) [hebrew]. *Cathedra* 16.
- Fagyal, Zsuzsanna. 2003. La prosodie du français populaire des jeunes. *Le Français aujourd'hui* 45–55.
- Fisherman, Haya. 2011. Talk given in the university of haifa, 27.12.11. URL <https://hazatetet.wordpress.com/2012/01/03/%D7%94%D7%9D-%D7%9E%D7%A9%D7%9C%D7%99%D7%9E%D7%99%D7%9D-%D7%90%D7%AA-%D7%94%D7%A7%D7%98%D7%99%D7%A2%D7%95%D7%AA-%D7%94%D7%9C%D7%A9%D7%95%D7%A0%D7%99%D7%95%D7%AA-%D7%91%D7%90%D7%9E%D7%A6%D7%A2%D7%95/>.
- Fletcher, Janet, Esther Grabe, and Paul Warren. 2005. Intonational variation in four dialects of english: the high rising tune .
- Floccia, Caroline, Jeremy Goslin, Frédérique Girard, and Gabrielle Konopczynski. 2006. Does a regional accent perturb speech processing? *Journal of Experimental Psychology: Human Perception and Performance* 32:1276.
- Foroni, Francesco, and Ulrich Mayr. 2005. The power of a story: New, automatic associations from a single reading of a short scenario. *Psychonomic Bulletin & Review* 12:139–144.
- Fox, J. 2019. John fox and sanford weisberg. *An R companion to applied regression, 3rd ed: Sage Publications [Google Scholar]. Published online .*
- Gafter, Roey J. 2014. The most beautiful and correct hebrew: Authenticity, ethnic identity and linguistic variation in the greater tel aviv area. Doctoral Dissertation, Stanford University.
- Gafter, Roey J. 2016a. Pharyngeal beauty and depharyngealized geek. *Raciolinguistics: How Language Shapes Our Ideas About Race* 185.
- Gafter, Roey J. 2016b. What's a stigmatized variant doing in the word list? authenticity in reading styles and hebrew pharyngeals. *Journal of Sociolinguistics* 20:31–58.

- Gafer, Roey J. 2019. 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*, 232–241. Routledge.
- Gafer, Roey J. 2020. Trilling as a sociolinguistic variable: Ethnicity and variation in the hebrew dorsal fricatives. *Sociolinguistic Studies* 14:483–503.
- Goffman, Erving. 1963. *Embarrassment and social organization*. John Wiley & Sons Inc.
- Gorman, James. 1993. Like, uptalk? URL <https://www.nytimes.com/1993/08/15/magazine/on-language-like-uptalk.html>.
- Grainger, Jonathan, and Johannes Ziegler. 2011. A dual-route approach to orthographic processing. *Frontiers in Psychology* 2:54. URL <https://www.frontiersin.org/article/10.3389/fpsyg.2011.00054>.
- Greenwald, Anthony G, and Shelly D Farnham. 2000. Using the implicit association test to measure self-esteem and self-concept. *Journal of personality and social psychology* 79:1022.
- Greenwald, Anthony G, Debbie E McGhee, and Jordan LK Schwartz. 1998. Measuring individual differences in implicit cognition: the implicit association test. *Journal of personality and social psychology* 74:1464.
- Greenwald, Anthony G, Brian A Nosek, and Mahzarin R Banaji. 2003. Understanding and using the implicit association test: I. an improved scoring algorithm. *Journal of personality and social psychology* 85:197.
- Gross, Nahum. 1976. Economic change in palestine at the end of the ottoman period [hebrew: tmurot kaklkalijot be-erets israel be-sof ha-tkufa ha-otmanit]. *Cathedra* 87–97.
- Hamdi-Sultan, Rim, Melissa Barkat-Defradas, Emmanuel Ferragne, and François Pellegrino. 2004. Speech timing and rhythmic structure in arabic dialects: a comparison of two approaches. In *International Speech and Communication Association*, 1613–1616.
- Hanani, Abualsoud, Hanna Basha, Yasmeen Sharaf, and Stephen Taylor. 2015. Palestinian arabic regional accent recognition. In *2015 International Conference on Speech Technology and Human-Computer Dialogue (SpeD)*, 1–6.
- Haršav, Binyāmîn. 1993. *Language in time of revolution*. Univ of California Press.

- Hart, Betty, and Todd R Risley. 2003. The early catastrophe: The 30 million word gap by age 3. *American educator* 27:4–9.
- Hasson, Shlomo, and Amiram Gonen. 1974. Ethnic differences in residential mobility in immigrant housing estates of medium-sized towns in israel [hebrew: hevdelim adati'im be-shinuj mekom megurim: shikunej ha-olim be-shulej arim benonijot be-israel]. *Megamot [Hebrew]* 310–315.
- Henkin-Roitfarb, Roni. 2011. Hebrew and arabic in asymmetric contact in israel. *Lodz Papers in Pragmatics* 6:195–228.
- Henshke, Yehudit. 2013a. On the mizrahi sociolect in israel: A sociolexical consideration of the hebrew of israelis of north african origin. *Journal of Jewish Languages* 1:207–227.
- Henshke, Yehudit. 2013b. Peripheral hebrew: Between error and loan. *Language Studies* 14–15.
- Henshke, Yehudit. 2013c. Sara shilo's "no gnomes will appear": a linguistic analysis. *Hebrew Studies* 265–284.
- Horesh, Uri. 2015. Structural change in urban palestinian arabic induced by contact with modern hebrew. In *Semitic languages in contact*, 198–233. Brill.
- Irvine, Judith T, and Susan Gal. 2000. *Language ideology and linguistic differentiation*, volume 1, 402–434. Wiley Blackwell Malden, MA, USA.
- Izre'el, Shlomo, and Vered Silber-Varod. 2009. "'omer lenateax lenateax': Perception of prosodic groups in spoken hebrew" (in hebrew). *Hebrew Linguistics* 13–33.
- Jacobs, Neil G. 2005. *Yiddish: A linguistic introduction*. Cambridge University Press.
- Kark, Ruth. 1984a. Changing patterns of landownership in nineteenth-century palestine: the european influence. *Journal of Historical Geography* 10:357–384.
- Kark, Ruth. 1984b. *Jaffa: the growth of a city, 1799-1917 (jafotsmixata shel 'ir [hebrew])*. Sifrija le-toldot ha-jishuv ha-jehudi be-erets Israel (the library of Jewish settlement in Israel).
- Karpinski, Andrew, and Ross B Steinman. 2006. The single category implicit association test as a measure of implicit social cognition. *Journal of personality and social psychology* 91:16.

- Khamis-Dakwar, Reem, Karen Froud, and Peter Gordon. 2012. Acquiring diglossia: mutual influences of formal and colloquial arabic on children's grammaticality judgments. *Journal of child language* 39:61–89.
- Khazoom, Aziza. 1999. Western culture, stigma, and social closure; the origins of ethnic inequality among jews in israel [hebrew]. *Sociologia Israelit [Israeli Sociology]* 385–428.
- King, Sharese, and Meghan Sumner. 2014. Voices and variants: Effects of voice on the form-based processing of words with different phonological variants. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 36.
- Kiparsky, Paul. 1968. How abstract is phonology? indiana university linguistics club. published 1973 as phonological representations. *Three Dimensions of Linguistic Theory*, ed. O. Fujimura, 1-136. Tokyo: TEC Co .
- Kraljic, Tanya, and Arthur G Samuel. 2006. Generalization in perceptual learning for speech. *Psychonomic bulletin & review* 13:262–268.
- Kraljic, Tanya, Arthur G Samuel, and Susan E Brennan. 2008. First impressions and last resorts: How listeners adjust to speaker variability. *Psychological science* 19:332–338.
- Kuznetsova, Alexandra, Per B Brockhoff, and Rune Haubo Bojesen Christensen. 2017. lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software* 82.
- Labov, William. 1972. *Sociolinguistic patterns*. 4. University of Pennsylvania Press.
- Labov, William, Sharon Ash, and Charles Boberg. 2008. *The atlas of north american english: Phonetics, phonology and sound change*. Walter de Gruyter.
- Lambert, Wallace E, Moshe Anisfeld, and Grace Yeni-Komshian. 1965. Evaluation reactions of jewish and arab adolescents to dialect and language variations. *Journal of personality and social psychology* 2:84.
- Leemann, Adrian, Marie-José Kolly, Francis Nolan, and Yang Li. 2018. The role of segments and prosody in the identification of a speaker's dialect. *Journal of Phonetics* 68:69–84.
- Leemann, Adrian, and Beat Siebenhaar. 2008. Perception of dialectal prosody. In *Ninth Annual Conference of the International Speech Communication Association*.
- Lefkowitz, Daniel. 2004. *Words and stones: The politics of language and identity in israel*, volume 26. Oxford University Press on Demand.



- Lenth, Russell, Henrik Singmann, Jonathon Love, Paul Buerkner, and Maxime Herve. 2018. Emmeans: estimated marginal means, aka least-squares means. r package vers 1.2. 4.
- Lev, Tali. 2008. "we shall erase the past of those who have a past": the protocol of the meeting of pm golda meir with the israeli black panthers. *Teoria ve-Bikoret "Theory and Critique"* .
- Lev-Ari, Shiri. 2016. Selective grammatical convergence: Learning from desirable speakers. *Discourse Processes* 53:657–674.
- Lev-Ari, Shiri, Robin Dodsworth, Jeff Mielke, and Sharon Peperkamp. 2019. The different roles of expectations in phonetic and lexical processing. In *INTERSPEECH*, 2305–2309.
- Lev-Ari, Shiri, and Sharon Peperkamp. 2014. An experimental study of the role of social factors in language change: The case of loanword adaptations. *Laboratory Phonology* 5:379–401.
- Levon, Erez, and Roey Gafter. 2019. “this is not europe”: Sexuality, ethnicity and the (re) enactment of israeli authenticity. *Discourse, Context & Media* 30:100287.
- Lewin-Epstein, Noah, and Yinon Cohen. 2019. Ethnic origin and identity in the jewish population of israel. *Journal of Ethnic and Migration Studies* 45:2118–2137.
- Lötzsch, Ronald. 1990. *Jiddisches wörterbuch*. Bibliogr. Institution.
- Matras, Yaron, and Leora Schiff. 2005. Spoken israeli hebrew revisited: Structures and variation. *Studia Semitica: The Journal of Semitic Studies Jubilee Volume* 145–191.
- Maye, Jessica, Richard N Aslin, and Michael K Tanenhaus. 2008. The weckud wetch of the wast: Lexical adaptation to a novel accent. *Cognitive Science* 32:543–562.
- Mehager, Tom. 2018. The new representation of the mizrachi: a sort of poke in the eye. URL <https://www.haokets.org/2018/06/01/%D7%94%D7%99%D7%99%D7%A6%D7%95%D7%92-%D7%94%D7%9E%D7%96%D7%A8%D7%97%D7%99-%D7%94%D7%97%D7%93%D7%A9-%D7%9E%D7%A2%D7%99%D7%9F-%D7%90%D7%A6%D7%91%D7%A2-%D7%91%D7%A2%D7%99%D7%9F/>.
- Melnik, Gerda Ana, and Sharon Peperkamp. 2019. Perceptual deletion and asymmetric lexical access in second language learners. *The Journal of the Acoustical Society of America* 145:EL13–EL18.

- Mizrachi, Nissim, and Hanna Herzog. 2012. Participatory destigmatization strategies among palestinian citizens, ethiopian jews and mizrahi jews in israel. *Ethnic and Racial Studies* 35:418–435.
- Monterescu, Daniel. 2019. Jaffa shared and shattered: Contrived coexistence in israel/palestine. *Mediterranean Historical Review* .
- Montgomery, Chris, and Emma Moore. 2018. Evaluating s (c) illy voices: The effects of salience, stereotypes, and co-present language variables on real-time reactions to regional speech. *Language* 94:629–661.
- Mor, Uri, and Ivy Sichel. 2015. Native hebrew and ha-gashash ha-chiver trio. *Carmillim* 133–182.
- Morag, Shelomo. 1963. *Hebrew language tradition of the yemenite jews (vol. 4)*.
- Morag, Shelomo. 1990. Ha'ivrit ha-xadasha be-hitgabshuta: Lashon be'aspakariya shel xevra (consolidation of modern hebrew: A language in the light of society). *Katedra* 70-92 .
- Nahir, Moshe. 1998. Micro language planning and the revival of hebrew: A schematic framework. *Language in society* 335–357.
- Nespor, Marina, Marcela Peña, and Jacques Mehler. 2003. On the different roles of vowels and consonants in speech processing and language acquisition. *Lingue e linguaggio* 2:203–230.
- Neuman, Yishai. 2012. The origin of the diphthong /ej/ in modern hebrew and the factors behind its distribution. "moca'a sel hatnu'a ey be-ivrit bat zmanenu vehagormim hamatnim et tfucata" [hebrew]. *Ha'ivrit safa haya* 237–260.
- Ozerov, Pavel. 2013. Intonation: Israeli hebrew 113:327–335.
- Ozerov, Pavel. 2019. This is not an interrogative: the prosody of “wh-questions” in hebrew and the sources of their questioning and rhetorical interpretations. *Language Sciences* 72:13–35.
- Pallier, Christophe, Angels Colomé, and Núria Sebastián-Gallés. 2001. The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science* 12:445–449.
- Peirce, Jonathan, and Michael MacAskill. 2018. *Building experiments in psychopy*. Sage.

- Peleg, Yaron. 2018. Conclusion: The legacy of hebrew orientalism. In *Orientalism and the hebrew imagination*, 126–142. Cornell University Press.
- Pharao, Nicolai, Marie Maegaard, Janus Spindler Møller, and Tore Kristiansen. 2014. Indexical meanings of [s+] among copenhagen youth: Social perception of a phonetic variant in different prosodic contexts. *Language in Society* 43:1–31.
- Pike, Kenneth L. 1945. The intonation of american english. .
- Powell, Michael JD. 2009. The bobyqa algorithm for bound constrained optimization without derivatives. *Cambridge NA Report NA2009/06, University of Cambridge, Cambridge* 26–46.
- Prechtel, Christine, and Cynthia G Clopper. 2016. Uptalk in midwestern american english. In *Speech Prosody*, 133–137.
- Purnell, Thomas, William Idsardi, and John Baugh. 1999. Perceptual and phonetic experiments on american english dialect identification. *Journal of language and social psychology* 18:10–30.
- 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, Franck, and Jacques Mehler. 1999. Language identification with suprasegmental cues: A study based on speech resynthesis. *The Journal of the Acoustical Society of America* 105:512–521.
- Ramus, Franck, Marina Nespør, and Jacques Mehler. 1999. Correlates of linguistic rhythm in the speech signal. *Cognition* 73:265–292.
- Ravid, Dorit Diskin. 1995. *Language change in child and adult hebrew: A psycholinguistic perspective*. Oxford University Press on Demand.
- Razi, Tami. 2011. Arab-jews? ethnicity, nationality and gender in mandate tel aviv / arab-jews? ethnicity, nationality and gender in mandate tel aviv. *Teoria ve-Bikoret "Theory and Critique"* 137–160.
- Ribke, Nahuel. 2004. “the ugly” vs. “the beautiful” israeli struggles over defining israeli identity. *Israel: Tel Aviv University. MA thesis, at <http://www.tau.ac.il/tarbut/tezot/Ribke/Ribke-MA.pdf>(accessed 4 March 2007),(in Hebrew, including English summary)* .

- Ritchart, Amanda, and Amalia Arvaniti. 2014. The form and use of uptalk in southern californian english. Dublin:[Sn].
- Roby, Bryan K. 2015. *The mizrahi era of rebellion: Israel's forgotten civil rights struggle 1948-1966*. Syracuse University Press.
- Rotbard, Sharon, and Orit Gat. 2015. *White city, black city: Architecture and war in tel aviv and jaffa*. Pluto Press London.
- Said, Edward. 1978. Orientalism: Western concepts of the orient. *New York: Pantheon* .
- Sando, Yves Talla. 2009. Upspeak across canadian english accents: Acoustic and sociopho-  
netic evidence. In *Proceedings of the 2009 annual conference of the Canadian Linguistic Association*.
- Sasson-Levy, Orna. 2013. A different kind of whiteness: Marking and unmarking of social boundaries in the construction of hegemonic ethnicity. In *Sociological Forum*, volume 28, 27–50. Wiley Online Library.
- Schilling-Estes, Natalie. 2008. Stylistic variation and the sociolinguistic interview: A re-consideration. *Rafael Monroy & Aquilino Sánchez (eds.)* 25.
- Schwarzwald, Ora. 1981. *Grammar and reality of the hebrew verb (dikduk u-meciut ba-po'al ha-ivri)*. Universitat Bar-Ilan.
- Sharon, Aryeh. 1951. *Principles of physical planning for israel (jesodot la-tixnun ha-fizi be-israel)*, 5–11. Government publishing [Hebrew].
- Shavit, Ya'akov. 1984. *From ivri to kna'ani (hebrew: me-ivri le-kna'ani)*. Domino and Tel Aviv University[Hebrew].
- Sheleg, Yair. 2010. *From old "ivri" to new jew: the renaissance of judaism in israeli society [hebrew]*. Israel Democracy Institute.
- Shohat, Ella. 2001. Rupture and return: The shaping of a mizrahi epistemology. *HAGAR: International Social Science Review* 2:61–92.
- Shohat, Ella. 2015. The question of judeo-arabic. *The Arab Studies Journal* 23:14–76.
- Shokeir, Vanessa. 2008. Evidence for the stable use of uptalk in south ontario english. *University of Pennsylvania Working Papers in Linguistics* 14:4.

- Strand, Elizabeth A. 1999. Uncovering the role of gender stereotypes in speech perception. *Journal of language and social psychology* 18:86–100.
- Sullivan, Gail M, and Anthony R Artino Jr. 2013. Analyzing and interpreting data from likert-type scales. *Journal of graduate medical education* 5:541–542.
- Sumner, Meghan, and Reiko Kataoka. 2013. Effects of phonetically-cued talker variation on semantic encoding. *The Journal of the Acoustical Society of America* 134:EL485–EL491.
- Sumner, Meghan, Seung Kyung Kim, Ed King, and Kevin B McGowan. 2014. The socially weighted encoding of spoken words: A dual-route approach to speech perception. *Frontiers in psychology* 4:1015.
- Sumner, Meghan, and Arthur G Samuel. 2009. The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language* 60:487–501.
- Svetozarova, Natalia, Yuri Kleiner, Tjeerd De Graaf, and Rogier Nieuweboer. 1999. Russian-yiddish: Phonetic aspects of language interference. In *Proc. of the XIV International Congress of Phonetic Sciences. San-Francisco*.
- Swirski, Shlomo. 1991. Israel: the district of separated tracks. (israel: mehoz ha-maslulim hanifradim [hebrew]). *Megamot [Hebrew]* 266–285.
- Swirski, Shlomo. 2018. Acting against the gapped system of education (lif'ol be-nigud le-maarexet ha-hinux ha-mefo'eret [hebrew]). URL <https://adva.org/he/education-against-inequality/>.
- Trudgill, Peter. 1972. Sex, covert prestige and linguistic change in the urban british english of norwich. *Language in society* 1:179–195.
- Turnbull, Rory, and Sharon Peperkamp. 2017. The asymmetric contribution of consonants and vowels to phonological similarity: Evidence from lexical priming. *The Mental Lexicon* 12:404–430.
- Tzfadia, Erez, and Oren Yiftachel. 2008. Sfariferia: the development towns and the mizrahi space (hebrew: Sfariferia: Arej ha-pituax ve-hamakom ha-mizrahi) 22–30.
- Vaughn, Charlotte, and Tyler Kendall. 2018. Listener sensitivity to probabilistic conditioning of sociolinguistic variables: The case of (ing). *Journal of Memory and Language* 103:58–73.

- Vicenik, Chad, and Megha Sundara. 2013. The role of intonation in language and dialect discrimination by adults. *Journal of Phonetics* 41:297–306.
- Walker, Abby, and Jen Hay. 2011. Congruence between ‘word age’ and ‘voice age’ facilitates lexical access. *Laboratory Phonology* 2:219–237.
- Weatherhead, Drew, and Katherine S White. 2018. And then i saw her race: Race-based expectations affect infants’ word processing. *Cognition* 177:87–97.
- Weber, Andrea, and Anne Cutler. 2004. Lexical competition in non-native spoken-word recognition. *Journal of memory and language* 50:1–25.
- Weirich, Melanie, Stefanie Jannedy, and Gediminas Schüppenhauer. 2020. The social meaning of contextualized sibilant alternations in berlin german. *Frontiers in Psychology* 11:2664.
- Xu, Kaiyuan, Brian Nosek, and Anthony Greenwald. 2014. Psychology data from the race implicit association test on the project implicit demo website. *Journal of Open Psychology Data* 2.
- Yaeger-Dror, Malcah. 1988. The influence of changing group vitality on convergence toward a dominant linguistic norm: An israeli example. *Language & Communication* 8:285–305.
- Yaeger-Dror, Malcah. 1993. Linguistic analysis of dialect “correction” and its interaction with cognitive salience. *Language Variation and Change* 5:189–224.
- Yekutieli-Cohen, Edna. 2010. Akiva aryeh weiss and the first hebrew city (akiva aryeh weiss ve-ha’ir ha-ivrit ha-rishona [hebrew]). *Cathedra* 131–148.
- Yiftachel, Oren. 2021. *Landed power: Israel/palestine between ethnocracy and creeping apartheid*. Resling [Hebrew].
- Zameret, Zvi. 2012. *The melting pot in israel: The commission of inquiry concerning the education of immigrant children during the early years of the state*. SUNY Press.
- Ziegler, Johannes C, Ludovic Ferrand, and Marie Montant. 2004. Visual phonology: The effects of orthographic consistency on different auditory word recognition tasks. *Memory & Cognition* 32:732–741.
- Zlotnick, Elad, Andrew. J. Dzikiewicz, and Yoav Bar-Anan. 2015. Minno.js (version 0.3). [Computer software] .

Zuckermann, Ghil'ad. 2005. Abba, why was professor higgins trying to teach eliza to speak like our cleaning lady?: Mizrahim, ashkenazim, prescriptivism and the real sounds of the israeli language. *Australian Association of Jewish Studies* .

Zuckermann, Ghil'ad. 2020. *Revivalistics: from the genesis of israeli to language reclamation in australia and beyond*. Oxford University Press.

Zuur, Alain F, Elena N Ieno, and Chris S Elphick. 2010. A protocol for data exploration to avoid common statistical problems. *Methods in ecology and evolution* 1:3–14.

## Appendix A

### Screenshots from an online debate on “Mizrahi speech”

This appendix includes screenshots of an online debate regarding an animated sketch show for Hanuka, a holiday whose traditional food is donuts. In the sketch, a woman is in a holiday party at work, and she decides to eat a donut and thus break her diet. The sketch follows the donut into the woman’s stomach, where a polite enzyme-man explains to the donut-woman about his new plan to use “empty carbs”, like herself, in a more efficient way. Another donut follows, and they tell the enzyme that they both mean to go directly to the buttocks. “Our family is there, our friends are there - we’re going to the buttocks”, they explain. They are joined by an avalanche of more donuts, wearing yellow-black football scarfs and hats, yelling a fan song with altered words, one of them an often-used loanword from Arabic: "jallah, to the ass!".

Some participants in the discourse perceive the donut characters to sound “Mizrahi”; others insist that there is no difference in the way these characters are portrayed compared with the enzyme. The screenshots in A.1, A.2 and A.3 are from a feminist and Mizrahi-activist Facebook pages, respectively. They raise concerns of body shaming on the one hand and racism on the other. In A.4 and A.4 are examples of participants who did not hear a difference between accents, and thought the sketch was funny. In A.6, a participant claims that not every usage of an accent in a sketch is racist. A.7 and A.8 agree with the page’s post that the sketch is covertly bigoted.





פוליטיקלי קוראת

December 13, 2017 · 🌐

Tammy Gottlieb כותבת לנו על הסרטון החדש של היוצרים ניר וגלי לחנוכה: זה הסרט המצחיק שרץ בחנוכה הזה. וברור שהוא כזה כי ניר וגלי תמיד הולך. וגם לצחוק על שמנות וגם להיות "קצת" גזענים. אני לא יכולה להתאפק מלהיות (שוב) זו שהורסת את המסיבה. אבל הסרטון הזה לא הצחיק אותי. סופגניה, שהיא פחמימה ריקה, מעוניינת להיכנס במערכת העיכול ישר לתחת ופוגשת בנציג השירות של הקיבה שמנסה לייעד אותה למקומות נכונים יותר ופחות מזיקים. בסדר. קצת מצחיק. והוא מסביר ומסביר, למה הכי טוב לה לא ללכת ישר לתחת. "מעכשיו גם פחמימות פשוטות יותר יתפרקו, יתנו אנרגיה זמינה יותר בגוף, וכך נשמור על סביבה אסתטית יותר, בריאה יותר וגוף שכיף לחיות בו לכולנו". והופ- אימג' של דמות שמנה מבואסת לעומת דמות רזה מאושרת. המוטיבציה לתזונה נכונה מכה שנית. לא נמאס לנו? לחיות בעולם שכל המוטיבציה שמושכת אותנו להתקיים היא המלחמה בכמה התחת שלנו גדול? לא למדנו שהמסרים האלה ממשיכים לשמר אידאלים שאנחנו לא יכולות ויכולים לעמוד בהם ומעודדים הפרעות אכילה שנגמרות במוות. אנחנו צריכות וצריכים לשאוף לאכול נכון כי זה בריא. לגרום לנו לתעב את עצינו כי הפחמימה הריקה שאכלנו הלכה לתחת זה כל כך הומור ישן, שזה מעייף.

ועוד נקודה על הסרט הזה. ברור באיזו מבטא מדברות הסופגניות הטיפשות שהן "פחמימות ריקות" ובאיזה מבטא מדבר מנהל הקיבה המנסה למכור לה להיות יותר אחרית. גם הגזענות הזו משעממת.

יאללה, הולכת להוסיף לי פחמימה ריקה לתחת.

בתמונה: האנזים מושיק מסביר על חשיבות הבריאות.

קשת 12

ארץ נהדרת

ניר וגלי

לינק לסרטון המלא בתגובות.

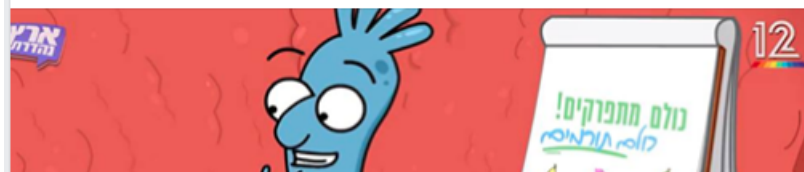


Figure A.1: Post in Politically Corret (lit 'political reader+fm'), a feminist facebook page


**הערסים באים**
...

13 בדצמבר 2017 · 13

**קצת גזענות בשם האור לחג**  
 ===

אז ניר וגלי, המצחיקנים הלאומיים התורנים, החליטו לשמח את עם ישראל בחג האור והעלו סרטון ויראלי מאין כמותו על הפיתוי שבאכילת סופגניות, השמנה וקמפיין תרומה לקהילה שהקיבה מנסה לרתום אליו את הסופגניות הזורמות אליה בהמוניהן.

אבל, ניר וגלי, למה אתם חושבים שגזענות זה מצחיק?

למה הסופגניות מדברות כמו מזרחיות, בעלות סממנים שמתויגים כ"פרחות" (מילה יפה שהגזענות הישראלית השמידה), הן פשוטות ורק רוצות להיות עם המשפחה שלהן, לא מוכנות להרתם למטרות הפוליטיות וה"ערכיות" של תרומה לקהילה?

למה מושיק, הפעיל החברתי-סביבתי, מדבר במילים גבוהות, מתנשא, בעל הערכים, המורכב והמתוחכם, שחושב לא רק על עצמו אלא על הכלל? ולמה חבורת הסופגניות שבסוף לובשת צעיפי בית"ר ירושלים וצועקת כביכול בהתלהמות? למה צעקות "העם דורש צדק חברתי" נתפסות כערכיות ואילו רצון הסופגניות להיות בקהילה שלהן ועם המשפחה נתפס כהמוני וברברי?

זה האבסורד שבתכניות ה"סאטירה" הישראליות מגב האומה ועד ארץ נהדרת. הן מגחכות ציבורים שלמים בשם נאורות וליברליות מדומיינות ובעצם חושפות את זרם התודעה המתנשא והגזעני של יוצריהן, שלאחרונה גם נחשף שלכאורה הם לא דווקא המעסיקים הכי נאורים. בכל זאת, אם שילמת למישהו תוכל לנהוג כלפיו בכל דרך.

Figure A.2: Post in Ha-'arsim ba'im (lit 'the Arsim are coming', Ars is a derogatory term depicting a Mizrahi man), a Mizrahi activists Facebook page

יש בכך רצון אמיתי לשינוי לטובת \*כלל\* הקהילה?  
הגיע הזמן לשנות את ההנהגה הישראלית, אבל  
קודם כל להחליף את ההנהגה של השמאל.  
וזו כל התורה של הערסים באים על רגל אחת, בונים  
שמאל אמיתי.

ניר וגלי  
ארץ נהדרת  
קשת 12 See Less

Figure A.3: Continuation of the post in Ha-'arsim ba'im from A.2

 **Linoy Shilian**

טרחני ממש.  
הסרטון הזה לגמרי עוסק ביום של רובנו, בטח בתקופת החג. כולנו  
מרגישים פחות טוב עם עצמנו כשאנחנו מכניסים לגוף שלנו ג'אנק (כן,  
גם גברים) ולא, זה לא רק בגלל השמנה, שמעתן על לחץ דם? מחלות  
לב? חלאס עם הפמניזם בשקל הזה, אין שום דבר רע בלצחוק קצת על  
חולשה שיש לכולנו\*, לא רק לנשים.  
ואגב, הסופגניות הן ערסיות/פאקצות ועד כמה שאני יודעת כאלו יש מכל  
עדה, זו בחירה שלכן לקטלג אותן בעדה מסויימת, זה נוח לכן, אותו דבר  
גם לגביי מנהל הקיבה- על איזה מבטא אתן מדברות בדיוק?

3y · Edited  10

Figure A.4: "There is no accent"



Figure A.5: “No body shaming and no accent”



Figure A.6: “Using an accent is not necessarily racist”



Figure A.7: “This bothered me as well, the sketches made by these cartoonists are always racist”



Figure A.8: “Thank you for bringing this up, the racism and body shaming bothered me as well”

## Appendix B

### Full models of Experiments 2-3

Analyses were conducted using the lmerTest package (Kuznetsova et al. (2017)) and the emmeans package (Lenth et al. (2018)) in the R software environment (R Core Team (2017)). A logistic model was used for accuracy/acceptance data, and a linear model for RT data.

**Experiment 2:** The models included random intercepts for participants and items, and Condition ([/h/] vs. [χ] items), Group (Pharyngeal vs. Merged linguistic background) and their interactions as fixed effects. The full models are presented in Tables B.1 and B.2 below.

**Experiment 3:** The models included random intercepts for participants and items, and the fixed effects Voice (Merged/ Pharyngealizer, sum-encoded), Condition (treatment contrast with 5 levels: No [/h]/[χ] (Baseline), [/h/] items, [χ] items, Inconsistent and Pathological, the latter two being relevant only for the Pharyngealizer voice blocks) and Group (sum-encoded comparisons between Ashkenazi vs. both generations of Mizrahi participants, and 1st vs. 2nd generation Mizrahi participants), including their interactions, as well as two additional fixed effects for which we did not consider interactions: Running mode (Offline/Online) and age. The bobyqa optimizer was used to allow the models to converge (Powell 2009). The full models are presented in Tables B.3 and B.4 below.

	Estimate	Std. Error	z-value	Pr(> z )
(Intercept)	3.63	0.39	9.37	< .0001 ***
Pharyngeal listener	1.33	0.64	2.08	0.04 *
Condition[/h/]	-1.13	0.27	-4.2	< .0001 ***
Pharyngeal listener:Condition[/h/]	0.02	0.39	0.05	0.96

**Table B.1: Accuracy model of Experiment 2. Number of participants: 70; Items: 34; Observations: 2380.**

	Estimate	Std. Error	Df	t-value	Pr(> t )
(Intercept)	7.16	0.04	99.95	184.71	< <b>0.0001</b> ***
Pharyngeal listener	-0.11	0.05	76.96	-2.08	<b>0.04</b> *
Condition[/h/]	0.08	0.03	38.53	2.34	<b>0.02</b> *
Pharyngeal listener:Condition[/h/]	0.03	0.02	2250.69	1.27	0.2

**Table B.2: Reaction times model of Experiment 2. Number of participants: 70; Items: 34; Observations: 2354.**

	Estimate	Std. Error	z-value	Pr(> z )
Intercept	4.29	0.24	17.82	< .0001 ***
Condition/γ/	-0.17	0.29	-0.6	0.55
Condition[/h/]	0.18	0.34	0.538	0.59
Condition[Inconsistent]	-0.84	0.32	-2.66	0.008 **
Condition[Pathological]	-3.17	0.28	-11.44	< .0001 ***
Voice[Merged]	0.19	0.18	1.047	0.29
Ash vs. Miz	0.05	0.11	0.42	0.67
1st vs. 2nd gen	0.07	0.22	0.34	0.73
Running[online]	0.04	0.31	0.126	0.9
Condition/γ/:Ash vs. Miz	0.2	0.14	1.4	0.16
Condition[/h/]:Ash vs. Miz	0.28	0.17	1.64	0.1
Condition[Inconsistent]:Ash vs. Miz	-0.44	0.17	-2.6	0.009 **
Condition[Pathological]:Ash vs. Miz	-1.04	0.13	-8.13	< 0.0001 ***
Condition/γ/:1st vs. 2nd gen	0.2	0.31	0.65	0.51
Condition[/h/]:1st vs. 2nd gen	-0.1	0.41	-0.24	0.81
Condition[Inconsistent]:1st vs. 2nd gen	-0.61	0.27	-2.29	0.02 *
Condition[Pathological]:1st vs. 2nd gen	-1.38	0.21	-6.51	< 0.0001 ***
Voice[Merged]:Ash vs. Miz	-0.05	0.09	-0.59	0.56
Voice[Merged]:1st vs. 2nd gen	-0.26	0.16	-1.6	0.11
Condition/γ/:Voice[Merged]:Ash vs. Miz	0.29	0.14	2.03	0.04 *
Condition[/h/]:Voice[Merged]:Ash vs. Miz	0.25	0.17	1.47	0.14
Condition/γ/:Voice[Merged]:1st vs. 2nd gen	-0.17	0.31	-0.555	0.58
Condition[/h/]:Voice[Merged]:1st vs. 2nd gen	0.35	0.41	0.85	0.39

**Table B.3: Acceptance rates model of Experiment 3. Number of participants: 75; Observations: 8325.**

	Estimate	Std. Error	Df	t-value	Pr(> t )
Intercept	6.861	0.04	86.95	175.99	< .0001 ***
Condition/ $\gamma$ /	0.03	0.01	176.6	2.060	0.041 *
Condition/ $/h/$ /	0.057	0.015	176.2	3.708	0.0003 ***
Condition[Inconsistent]	0.032	0.02	200.5	1.580	0.11570
Condition[Pathological]	0.137	0.022	281.9	6.215	< .0001 ***
Voice[Merged]	-0.025	0.011	147.7	-2.281	0.024 *
Ash vs. Miz	0.008	0.009	82.62	0.899	0.371
1st vs. 2nd gen	-0.022	0.02	80.11	-1.136	0.259
Running[online]	0.056	0.029	75.03	1.950	0.055 .
Age	0.002	0.0009	74.9	1.953	0.054 .
Condition/ $\gamma$ /:Voice[Merged]	0.022	0.01	176.6	1.417	0.158
Condition/ $/h/$ :Voice[Merged]	0.013	0.015	176.2	0.859	0.391
Condition [ $\gamma$ ]:Ash vs. Miz	0.002	0.004	7631	0.349	0.727
Condition/ $/h/$ :Ash vs. Miz	-0.011	0.004	7633	-2.501	0.012 *
Condition[Inconsistent]:Ash vs. Miz	0.018	0.007	7645	2.711	0.007 **
Condition[Pathological]:Ash vs. Miz	0.033	0.008	7702	4.057	< .0001 ***
Condition/ $\gamma$ :1st vs. 2nd gen	-0.005	0.008	7630	-0.612	0.54
Condition/ $/h/$ :1st vs. 2nd gen	0.008	0.008	7624	1.026	0.305
Condition[Inconsistent]:1st vs. 2nd gen	0.011	0.012	7636	0.888	0.375
Condition[Pathological]:1st vs. 2nd gen	0.046	0.018	7707	2.583	0.01 **
Voice[Merged]:Ash vs. Miz	0.003	0.002	7612	1.423	0.155
Voice[Merged]:1st vs. 2nd gen	-0.003	0.004	7611	-0.658	0.51
Condition [ $\gamma$ ]:Voice[Merged]:Ash vs. Miz	0.002	0.004	7627	0.505	0.61351
Condition/ $/h/$ :Voice[Merged]:Ash vs. Miz	-0.003	0.004	763	-0.786	0.43167
Condition/ $\gamma$ :Voice[Merged]:1st vs. 2nd gen	0.008	0.008	7628	0.987	0.32391
Condition/ $/h/$ :Voice[Merged]:1st vs. 2nd gen	0.01	0.008	7621	2.283	0.022 *

**Table B.4: Reaction times model of Experiment 3. Number of participants: 75; Observations: 7877; files: 195.**



## Appendix C

### Full models of Experiments 4-5

**Experiment 4:** A merged set of data was constructed from the results of studies 4a and 4b, in order to assess the general success rates in categorization, based on the first two practice categorization blocks (blocks 1,2). In block 1, participants had to categorize [h] and [χ], and in block 2, they categorized [e] and [ej]. Therefore, it was not possible to compare the success rates in all four stimulus types within the same model. In order to compare the relative success rates, the following models were constructed:

- To assess whether the distinction between [e] and [ej] was easier to make than the distinction between [h] and [χ] (or vice versa), a model with the fixed effects of Categorization Block ([e]-[ej] or [h]-[χ]) and Guise (Broadcaster/Yemenite) was used, with the random intercepts Participant and Item.
- To assess whether one of the categories was more easily identified (e.g., whether participants responded to [ej] faster or more accurately than to the other category [e]), a model with a subset of the data (only block 1) was used, with the fixed effect Stimulus Type and Guise (Broadcaster/Yemenite), and the random intercepts Participant and Item. A model for block 2 was constructed with the same question in mind.
- For each type of analysis, a generalized linear regression was used for accuracy and a linear mixed effects regression for the (log-transformed) RTs.

	Estimate	Std. Error	Df	t-value	Pr(> t )
(Intercept)	7.29	0.04	65.12	179.23	$< 2e-16$ ***
categorization([h]~[x])	0.024	0.046	34.13	0.525	0.6029
Guise[Yemenite]	0.13	0.032	107	4.084	$8.57e-05$ ***
categorization([h]~[x]):Guise[Yemenite]	0.054	0.03	107	1.800	0.0746 .

**Table C.1: Reaction times model of the categorization (practice) blocks of Experiment 4: Log RT, as predicted from the categorization task ([h]~[x] vs. [ej]~[e]) and Guise (Yemenite vs. Broadcaster). Number of trials = 4360, 109 participants, 20 media items.**

	Estimate	Std. Error	z-value	Pr(> z )
(Intercept)	-2.52	0.21	-11.74	< 2e-16 ***
categorization([h]~[x])	0.45	0.28	1.6	0.11
Guise[Yemenite]	0.4	0.13	3.15	0.00163 **
categorization([h]~[x]):Guise[Yemenite]	-0.012	0.14	-0.085	0.93212

**Table C.2: Error rate model of the categorization (practice) blocks of Experiment 4: rate of errors, as predicted from the categorization task ([h]~[x] vs. [e]~[e]) and Guise (Yemenite vs. Broadcaster). Number of trials = 4360, 109 participants, 20 media items.**

**Experiment 5:** Only results from participants who had an overall success rate of above 75% were used in the analyses (the same dataset used for the d-score analysis). The fixed factors were Guise (Argentinian/Arab, sum-contrasted), Condition ([r]+[h] / [r]+[x], sum-contrasted), Order ([r]+[h] was presented first/second; sum-contrasted), and Stimulus ([r]/[h]/[x]; Helmert-contrasted, such that [r] trials were contrasted with both types of other trials and [h] and [x] trials were contrasted with each other). The Helmert contrast was chosen for the factor Stimulus two reasons: (i) [r]-items were different recordings across guises, and (ii) the task can be viewed as two sub-tasks: association of [r] to either of the other variants; and for [h]/[x] items, the same association task in addition to phonetic categorization. That is why [r] can be viewed as a sort of baseline, that involves only one of the requirements in order to perform the task well. Participant and Item were added as random intercepts.

	Estimate	Std. Error	z-value	Pr(> z )
(Intercept)	-2.71647	0.14	-19.410	< .0001 ***
Guise[Manuel]	-0.11127	0.22	-0.510	0.610390
Condition[h+r]	0.13	0.11	1.192	0.233259
Order[1st]	-0.04	0.11	-0.351	0.725941
stimulus[h] vs. [x]	-0.03	0.11	-0.271	0.786515
stimulus[x]+[h] vs. [r]	-0.61426	0.08	-7.67	< .0001 ***
Guise[Manuel]:Condition[h+r]	0.38	0.22482	1.678	0.093437 .
Guise[Manuel]:Order[1st]	-0.04	0.22	-0.185	0.85
Condition[h+r]:Order[1st]	0.57	0.44	1.3	0.19
Guise[Manuel]:stimulus[h] vs. [x]	0.16	0.08	2.06	0.04 *
Guise[Manuel]:stimulus[x]+[h] vs. [r]	0.18	0.1	1.74	0.08 .
Condition[h+r]:stimulus[h] vs. [x]	0.23	0.078	2.99	0.0028 **
Condition[h+r]:stimulus[x]+[h] vs. [r]	0.13	0.1	1.28	0.2
Order[1st]:stimulus[h] vs. [x]	0.26	0.077	3.360	0.0008 ***
Order[1st]:stimulus[x]+[h] vs. [r]	-0.4	0.1	-3.89	< .0001 ***
Guise[Manuel]:Condition[h+r]:Order[1st]	0.18	0.87	0.2	0.84
Guise[Manuel]:Condition[h+r]:stimulus[h] vs. [x]	0.3	0.155	1.955	0.050531 .
Guise[Manuel]:Condition[h+r]:stimulus[x]+[h] vs. [r]	0.57	0.21	2.79	0.005320 **
Guise[Manuel]:Order[1st]:stimulus[h] vs. [x]	-0.098	0.15	-0.634	0.526
Guise[Manuel]:Order[1st]:stimulus[x]+[h] vs. [r]	-0.187	0.21	-0.91	0.36
Condition[h+r]:Order[1st]:stimulus[h] vs. [x]	-0.54	0.155	-3.49	0.0005 ***
Condition[h+r]:Order[1st]:stimulus[x]+[h] vs. [r]	0.65	0.21	3.15	0.001651 **
Guise[Manuel]:Condition[h+r]:Order[1st]:stimulus[h] vs. [x]	0.132	0.31	0.425	0.671
Guise[Manuel]:Condition[h+r]:Order[1st]:stimulus[x]+[h] vs. [r]	-0.07	0.41	-0.178	0.859

**Table C.3: Summary statistics of the generalized linear mixed model of Experiment 5 using the bobyqa optimizer, with accuracy as the dependent variable and the fixed factors Guise (Manuel/Waseem), Stimulus type ([h]/[x]/[r]), Condition ([h]+[r] vs. [x]+[r]) and Block Order (1st/2nd). Number of trials: 9668, 51 participants, 15 media items.**

	Estimate	Std. Error	Df	t-value	Pr(> t )
(Intercept)	7.06	0.025	56.72	276.492	< .0001 ***
stimulus[h] vs. [x]	0.04	0.016	12.06	2.34	0.04 *
stimulus[x]+[h] vs. [r]	-0.05	0.01	12.23	-5.23	0.0002 ***
Condition[h+r]	0.012	0.006	9585	1.853	0.064 .
Guise[Manuel]	-0.03	0.044	47.05	-0.694	0.491
Order[1st]	0.04	0.006	9585	6.204	< .0001 ***
stimulus[h] vs. [x]:Condition[h+r]	0.026	0.0076	9585	3.45	< .0001 ***
stimulus[x]+[h] vs. [r]:Condition[h+r]	-0.0006	0.0046	9585	-0.138	0.89
stimulus[h] vs. [x]:Guise[Manuel]	0.032	0.008	9585	4.202	< .0001 ***
stimulus[x]+[h] vs. [r]:Guise[Manuel]	0.023	0.0046	9585	5.148	< .0001 ***
Condition[h+r]:Guise[Manuel]	0.066	0.013	9585	5.203	< .0001 ***
stimulus[h] vs. [x]:Order[1st]	0.009	0.0076	9585	1.165	0.243892
stimulus[x]+[h] vs. [r]:Order[1st]	0.004	0.005	9585	0.809	0.418581
Condition[h+r]:Order[1st]	0.001	0.089	47.05	0.017	0.986811
Guise[Manuel]:Order[1st]	-0.001	0.013	9585	-0.097	0.922856
stimulus[h] vs. [x]:Condition[h+r]:Guise[Manuel]	0.038	0.015	9585	2.492	0.013 *
stimulus[x]+[h] vs. [r]:Condition[h+r]:Guise[Manuel]	0.017	0.009	9585	1.799	0.072029 .
stimulus[h] vs. [x]:Condition[h+r]:Order[1st]	-0.055	0.015	9585	-3.608	0.0003 ***
stimulus[x]+[h] vs. [r]:Condition[h+r]:Order[1st]	0.028	0.009	9585	3.058	0.002 **
stimulus[h] vs. [x]:Guise[Manuel]:Order[1st]	-0.018	0.015	9585	-1.214	0.225
stimulus[x]+[h] vs. [r]:Guise[Manuel]:Order[1st]	-0.0067	0.009	9585	-0.718	0.47
Condition[h+r]:Guise[Manuel]:Order[1st]	0.156	0.178	47.05	0.878	0.38
stimulus[h] vs. [x]:Condition[h+r]:Guise[Manuel]:Order[1st]	0.016	0.03	9585	0.53	0.596081
stimulus[x]+[h] vs. [r]:Condition[h+r]:Guise[Manuel]:Order[1st]	-0.01	0.018	9585	-0.556	0.578269

**Table C.4: Full Reaction times model of Experiment 5: Log RT, as predicted from the fixed effects Guise, Condition, Order and stimulus (including interactions), and participant and media as random effects. Number of trials: 9668, 51 participants, 15 media items.**

## תקציר

תזה זו בוחנת תפיסה של שונות (variation) פונולוגית מבוססת זהות אתנית: משמעותה החברתית עבור דוברי עברית, וההשפעות שלה על עיבוד שפה דבורה בזמן אמת. החלק הראשון מספק תמונת מצב עדכנית של השונות הפונולוגית בעברית מודרנית, בדגש על התפיסות השונות בספרות ובציבור הרחב לגבי שונות לשונית על רקע השסע האתני בין יהודים "אשכנזים" ו"מזרחים" בישראל. החלק השני עושה שימוש במשתנים פונולוגיים בעלי משמעות חברתית על מנת לחקור את המנגנונים הקוגניטיביים של עיבוד שפה דבורה, בדגש על האופן שבו זהות הדובר כפי שנתפשת על ידי הנמען משפיעה על זיהוי מילים דבורות. בפסקאות הבאות, אסכם את הממצאים תוך הדגשת התרומה של המחקר הנוכחי לחקר שונות לשונית בעברית מודרנית, לשאלות ותיקות ביחס לייצוג המנטלי של קטגוריות פונולוגיות ולמודלים של עיבוד מילים אודיטורי.

## ממצאים סוציו-לשוניים

מבחינה חברתית, המחקר הנוכחי מספק מסגרת תיאורית חדשה להבנת השימור של משתנים לשוניים המגיעים מדיאלקטים שונים של ערבית בקרב דוברי עברית מודרנית. הספרות המוקדמת על שונות לשונית מבוססת זהות אתנית בעברית מודרנית מדגישה תהליך של שחיקה מהירה ורוחבית בשונות הלשונית, והתכנסות לכיוון נורמה המובלת על ידי דוברים אשכנזים (למשל, Blanc 1968, Yaeger-Dror 1988). אני מציעה כי תהליך ההתכנסות המדובר אינו תקף באותה המידה עבור כלל האוכלוסייה הלא-אשכנזית, ויובן טוב יותר כסדרה של מקרים פרטיים, במסגרתם דוברת בוחרת להתאים את עצמה לנורמה לא-מסומנת בסביבה בעלת רוב אשכנזי. עקב פיזור האוכלוסייה בישראל, שמלכתחילה יצר מושבים, קיבוצים, שכונות ועיירות פיתוח הומוגניות מבחינה אתנית, דוברים מזרחים בפריפריה שמרו ופיתחו משתנים לשוניים שמקורם בשפות אמם, בעיקר ערבית. משתנים אלה, בהם לקסיקליים, תחביריים ופונולוגיים, עדיין נשמרים ומוסיפים להתפתח בחלקים מסוימים של ישראל, ומתוארים כוריאנט של עברית שלעתיים מכונה "מזרחית" (Matras and Schiff 2005) או "פריפריאלית" (Henshke 2013b). במקרים אחרים, המשתנים הללו מוסווים תחת קטגוריה חברתית אחרת ("מעמד סוציו-אקונומי נמוך").

טענתי המרכזית בפרק הראשון היא שמרבית הספרות הנוגעת לשונות לשונית בעברית מודרנית סובלת מהטיה עיקשת, שמביאה חוקרים להתעלם משונות, לדווח על התכנסות מבטאית חד כיוונית, ובמחקרים התפתחותיים לייחס את רוב השונות לגורמים סוציו-אקונומיים או אינדיבידואליים-קוגניטיביים במקום לרקע הלשוני של הדובר. את מקורה של ההטייה, לפחות בחלקה, ניתן לייחס לתהליך של איקוניזציה (Irvine and Gal 2000) של השפה כסמל לעלייתה מחדש של האומה העברית, תהליך שקשור בדוקטרינת כור ההיתוך שרווחה בעשורים המוקדמים לאחר הקמת מדינת ישראל. דיונים (מקצועיים-בלשניים ושאינם מקצועיים) מהתקופה בה המחקרים הראשונים על שונות בעברית מודרנית

יצאו לאור, כמו גם נתונים גיאוגרפיים והיסטוריים וחשוב מכל, תמונת המצב הנוכחית של של שונות בשפה, כולם סותרים את הטענה שהשונות בשפה נמצאת בתהליך שיטתי ורוחבי של דעיכה. כפי שהצעתי מעלה, חלק מהוריאנטים נעלמו בהדרגה בגלל סטריאוטיפיזציה ועקב זאת בחירה של הדוברים להעלימם, בזמן שוריאנטים אחרים משגשגים בקהילות בהן הסוציולקט המזרחי הוא הרווח. יש לציין שסגרגציה בין קהילות והומוגניות בתוכן הם באופן כללי תנאים המעודדים התפתחות, ולא הפחתה, של שונות לשונית. בהקשר הזה, נראה שגם דוברים מרקע אשכנזי מאזורים מסוימים במרכז הארץ, נוטים להיבדל בהפקה בדרכים חדשות ומזוהות.

עדויות אמפיריות ראשונות התומכות בטענה האחרונה מופיעות בפרק 2. בניסוי 1, דוברים ילידיים של עברית האזינו להקלטות קצרות (כחמש שניות כל אחת) של דוברי עברית מקבוצות גיל, מגדר ורקעים לשוניים מגוונים, כאשר הדגש היה על דגימה של הקלטות מסביבות לשוניות בעלות רוב אשכנזי או מזרחי. עבור כל הקלטה, המשתתפות/ים היו צריכות לדרג באיזו מידה הן בטוחות שהדובר/ת בהקלטה ממוצא אשכנזי או מזרחי, כאשר קצה אחד של הסקאלה סומן "בטוח שאשכנזי/ה", הקצה השני "בטוח שמזרחי/ת", ונקודת האמצע "אני לא בטוח/ה". בין הוריאנטים הלשוניים שנכללו בהקלטות היו חי"ת ועי"ן גרוניות ורי"ש מתגלגלת, המזוהות עם מזרחים (ראו למשל בן טולילה 1984), ודיפתונג בהגיית מילים כגון מימד, המזוהה עם אשכנזים (נוימן 2012). בקרב הצעירים שהוקלטו (מחצית מכלל ההקלטות), אף אחד לא השתמש במשתנים לשוניים הידועים מהספרות כמסומנים לזהות אתנית מסוימת. חרף זאת, לפי תוצאות הניסוי, חלק מדוברי/ות העברית הצעירים/ות נתפסו כשונים זה מזה מבחינה אתנית באופן שיטתי. התוצאות גם הדגימו ש"להישמע אשכנזי" אינה תמיד ברירת המחדל: בעוד שחלק מהדוברים דורגו כנשמעים מזרחים או אשכנזים, אחרים לא נפלו לאף קטגוריה. העדות הזו מחזקת את התפיסה שהוצגה בפרק 1, לפיה אף אחד מהסוציולקטים אינו צריך להיחשב לברירת מחדל שביחס אליה שופטים את האחר; "אשכנזי" אינה בהכרח זהות לא מסומנת.

ניתוח אקוסטי של ההקלטות ביחס לשיפוטי האתניות מצא קורלציה בין תפיסת אתניות ובין קצב, שנמדד כאחוז המשך של תנועות לעומת עיצורים בשטף הדיבור: ככל שתנועות היוו חלק קטן יותר מכלל הדיבור, שיפוטי האתניות נטו לכיוון יותר אשכנזי. התוצאה הזו מדגימה קורלציה בלבד - מבחינה אקוסטית יש בכלל הקלטה מספר לא ידוע של משתנים שלא נבחנו, וייתכן שקצב נמצא בקורלציה עם חלקם, כך שמשתנה לא ידוע אחר הוא שמשפיע על תפיסת המאזינים. עם זאת, קצב משחק תפקיד משמעותי בזיהוי דיאלקטים בשפות רבות. כך למשל, דוברי ערבית מרקעים דיאלקטיים מגוונים הצליחו להבחין בין דיאלקטים צפון-אפריקאים לבין דיאלקטים מזרחיים ב-98% מהמקרים, והיכולת של מאזינים לתפוש גם הבחנות דקות יותר, למשל בין טוניסאית למרוקאית, הושגה על אחוז התנועות מתוך כלל הדיבור (Barkat et al. 1999). גם דוברי גרמנית שוויצרית הבחינו בין ארבע קבוצות דיאלקטים בעיקר בהתבסס על קצב (Leemann & Siebner 2008; Leemann et al. 2018); וכאשר הושו הפקות של איטלקית לעומת ספרדית, במשפטים בהם הפריטים הלקסיקליים זהים והשונות הסגמנטלית מינימלית, מאזינים נסמכו על קצב יותר מעל כל מאפיין אחר בסיווג השפה (De Mareüil et al. 2006). בנוסף, בצרפתית של הפרברים, שמרבית דובריה ממוצא

צפון אפריקאי, נמצאה תופעה פרוזודית דומה לזו של הוריאנט המזרחי של עברית מבחינת קצב (Fagyal 2004), שיוחסה להשפעה של ערבית מערבית (מרוקו, טוניס, אלג'יר ולוב), דבר המרמז למקור משותף. כלומר, התוצאה הנוכחית מציעה נקודת פתיחה למחקר ממוקד יותר של קצב בעברית.

המסקנה המרכזית שעולה מהחלק הראשון של התזה (פרקים 1 ו 2) היא שהתפיסה לפיה שונות לשונית בעברית נמצאת בנסיגה מתמדת צריכה לפנות מקום לבחינה מדוקדקת של השונות הקיימת בשפה ושל התפיסות של דוברות עברית בנוגע אליה. מחקרים סוציו-בלשניים על עברית מודרנית יפיקו תועלת מתוצאות ניסוי 1, לפיו יש יותר מברירת מחזל אחת לשפה הדבורה. כך, ניתן יהיה לפנות לתעד את המאפיינים הלשוניים של קהילות דוברים הרואות את עצמן כמובחנות, בהתבסס על מאפיינים תרבותיים, אידאולוגיים, גיאוגרפיים ואתניים. חשוב מכל, חוקרות וחוקרים העוסקים בחינוך והתפתחות קוגניטיבית צריכים להיות מודעים לנורמות הפקה מגוונות בקרב הקהילות אותן הם חוקרים, במיוחד כאשר הם מעריכים את ההתפתחות התחברית והמורפולוגית של ילדים.

### ממצאים הנוגעים לתאוריה פונולוגית

החלק השני בתזה עוסק ביחסי הגומלין בין תפיסה חברתית ופונולוגית בזמן עיבוד מילים אודיטורי, באמצעות מתודולוגיה ניסויית כמותית. המחקרים בו משתמשים בוריאנטים לשוניים מסומנים ובפרסונות סטריאוטיפיות המפיקות אותם על מנת להעריך את הייצוגים הפונולוגיים של דוברים מרקעים לשוניים מגוונים מצד אחד, ואת התפקיד של עיבוד לשוני מבוסס-דובר בפענוח הקלט האקוסטי מן הצד השני.

פרק 3 (ניסויים 2-3) מתמקד בעיבוד ובייצוגים המנטליים של ההגה המסומן חברתית [h] (חי"ת) על ידי דוברים מרקע לשוני שבו חי"ת נהגית כחוכך לועי (דוברים מרקע "מזרחי", כלומר מבחין), לעומת דוברים מרקע לשוני שבו הגיית חי"ת מתמזגת עם הגיית כ"ף, החוכך/trill הוילוני/ענבלי (דוברים מרקע "אשכנזי", כלומר ממוזג). מרבית דוברי העברית הילידיים הצעירים ממזגים בין שני ההגאים בהפקה, כולל דוברים שהם דור שני להורים המבחינים בהגייה בין הקטגוריות. נשאלת השאלה האם הייצוגים המנטליים של מילים הכוללות את ההגה חי"ת עשויים להיות נבדלים בקרב דוברים בעלי אותה הפקה; כלומר, האם הקלט הלשוני בתקופת רכישת השפה ומאז השפיע על התפיסה של בנות ובני הדור השני לדוברים שמבחינים בין הקטגוריות, גם ללא כל השפעה ניכרת על ההפקה.

על מנת להעריך את המעמד של העיצור חי"ת בייצוג המנטלי של דוברים מרקעים שונים, נערכו שני ניסויים. בניסוי 2, דוברים מרקע ממוזג ומבחין האזינו להקלטות של מילות תפל הכוללות את חי"ת או כ"ף, דוגמת \*סעוכה או \*מחכנ. המשתתפים התבקשו לקטלג מהר ככל האפשר ובאופן מדויק ככל האפשר לגבי כל מילה האם היא נהגתה בחי"ת או

בכ"ף. נמצא כי דוברים מרקע ממוזג היו רק מעט פחות מדויקים ויותר איטיים בתגובותיהם ביחס לדוברים מרקע מבחין. כלומר, ברמה הפונטית, נראה שדוברים מכל הרקעים מסוגלים להבחין היטב בין הקטגוריות.

ניסוי 3 בחן את הייצוגים הלקסיקליים של פריטים הכוללים את הקטגוריות חיי"ת וכ"ף באמצעות מטלת החלטה לקסיקלית, במסגרתה בכל סיבוב יש לקבוע אם המילה שנשמעה אמיתית או מילת תפל, כאשר מחצית מהפריטים במטלה הם מילות תפל. על מנת להעריך האם חיי"ת וכ"ף נחשבות שקולות זו לזו, או נבדלות ברמת הייצוג, חלק מהמילים האמיתיות עברו מניפולציה במסגרתה חיי"ת וכ"ף הוחלפו זו בזו. בהחלפה של חיי"ת לכ"ף אין כל חדש מבחינת דוברות עברית: זהו כיוון המיזוג שנפוץ בקרב רוב הדוברים, כך שמילים כמו "מחר" ו"מכר" נהגות באותו האופן. לעומת זאת, כאשר כ"ף מוחלפת בחיי"ת, מתקבלת צורה פתולוגית שאינה נהגית באופן הזה על ידי אף דובר ילידי של עברית. אם דוברי עברית ממוזגים מתייחסים לשתי הקטגוריות כשקולות, תהליך העיבוד במסגרתו מילה כמו **מחוג** שנהגית בחיי"ת גרונית ממוזגת ללקסיקון הממוזג אינו שונה בעיקרון מתהליך שבו צורה פתולוגית כמו \***מחונית** תמופה ל**מכונית**. כלומר, הניבוי הוא שפריטים כאלה ייחשבו תקינים על ידי דוברים ממוזגים, כיוון שחיי"ת על פני השטח בכל מקרה ממוזגת לכ"ף ברמת הייצוג המנטלי. פרט לדוברים הממוזגים, השתתפו בניסוי שתי קבוצות נוספות: דוברים מבחינים (מזרחים מבוגרים, ההוגים את חיי"ת וכ"ף כקטגוריות נבדלות), ודור שני למבחינים (הוגים קטגוריה אחת ממוזגת אך נחשפו להבחנה בילדות). עבור הקבוצות הללו, נצפה לקושי עם הפריטים הפתולוגיים, שכן במערכת המבחינה חיי"ת על פני השטח ממוזגת לחיי"ת מנטלית, וכך גם כ"ף; כלומר, אין דרך למפות חיי"ת על פני השטח לכ"ף מנטלית.

תוצאות ניסוי 3 הראו שהדוברים מרקע מבחין, דור ראשון ושני, דחו בשיעור גבוה פריטים פתולוגיים דוגמת \***מחונית** בחיי"ת גרונית, וגם הגיבו אליהם בצורה איטית בהשוואה לקבוצת המאזינים הממוזגים. התוצאה הזו היא עדות לכך שהבדלים בייצוג המנטלי של הגאים בלקסיקון יכולים להתקיים גם בין דוברים שעל פני השטח הוגים את אותן צורות: דור שני למבחינים, שממוזגים בין הקטגוריות בהגייה בדיוק כמו דור שני לממוזגים, בכל זאת שומרים על ייצוגים מובחנים בין חיי"ת לכ"ף בלקסיקון. כפי שמחקרים על רכישת שפה שנייה בקרב דוברי אנגלית, הולנדית וצרפתית הראו בעבר, ישנה א-סימטריה מובנה בקידוד בין הקטגוריה הדומה פונטית לזו שבשפת האם שלהם (במקרה הזה החוכך הוילוני/ענבלי כ"ף), לבין הקטגוריה ה"חדשה", שנתפשת כתת-קטגוריה שלה (במקרה הזה חיי"ת, שנתפשת כסוג של כ"ף; Melnik & Peperkamp 2019, Darcy et al. 2013, Weber & Cutler 2004).

המסקנה לפיה דוברים ממוזגים לא מבחינים בין חיי"ת וכ"ף בלקסיקון גם משליכה על שאלה ותיקה בבלשנות הגנרטיבית. מאז דפוסי הצליל של אנגלית (The Sound Patterns of English, Chomsky & Halle 1968); או בקיצור (SPE), הספר שפרש לראשונה את המשנה שמאז נחשבת לתיאוריית החוקים הקלאסית של הפונולוגיה הגנרטיבית, השדה אמביוולנטי לגבי מידת האבסטרקציה האפשרית של ייצוגים מנטליים בפונולוגיה. ה-SPE איפשר



ייצוגים אבסטרקטיים מאוד. הייצוגים המנטליים (Underlying Representations; URs) המרכיבים את הלקסיקון עשויים להיבדל מייצוגי השטח (Phonetic Representations; PRs) במידה רבה, עד כדי כך שייצוגים מנטליים מסוימים לעולם לא יופיעו על פני השטח. קיפרסקי (Kiparsky 1968), בביקורת על המאפיין הזה של SPE, טען שרמת האבסטרקטיות הזו מעלה בעיות רציניות עבור רכישת שפה, והציע תנאי שימנע מערכות יחסים מסוימות בין הייצוגים המנטליים לייצוגי השטח:

**תנאי האלטרנציה:** ייצוג מנטלי חייב להופיע על פני השטח לפחות בסביבה פונולוגית אחת. אין אף אלמנט בייצוג המנטלי שלא מגיע לעולם לפני השטח.

תנאי האלטרנציה שולל מקרים שכונו בספרות "נטרול מוחלט" (absolute neutralization), כלומר מקרים בהם קטגוריה בייצוג המנטלי לא מגיעה אף פעם לייצוג השטח. עבור הדוברים הממוזגים שהשתתפו בניסויים 2-3, זהו בדיוק המקרה עם העיצור הלועי חי"ת: העיצור לעולם לא מגיע לפני השטח בדיאלקט אותו הם דוברים ואליו הם חשופים מרבית הזמן, אבל עדיין יש עדויות פונולוגיות סיסטמטיות להבחנה בין חי"ת לכ"ף, שתחת גישת SPE נובעים מהבדלים בייצוג המנטלי של הקטגוריות חי"ת וכ"ף. כך למשל, כ"ף יכולה להתחלף עם k בעמדות מסוימות (כיתב-יכתב kitev-jexatev) בעוד שחי"ת אטימולוגית לעולם לא תתחלף עם [k] (למשל, חיפש-יחפש xipes-jexapes); חי"ת גורמת להנמכת תנועה ל[a] וכ"ף לא (למשל, עורכת orexet לעומת אורחת oraxat). חומסקי והאלי (1968) הניחו שהבדלים שיטתיים כאלה בתפוצת העיצורים יכולים לבסס הבחנה קטגורית ביניהם, גם בהיעדר עדות פונטית. לשיטתו של קיפרסקי, זה בלתי אפשרי. ניסוי 3 תומך בתנאי האלטרנציה: דוברים ממוזגים שלא נחשפו לעדות פונטית לנבדלות בין הקטגוריות לא הראו עדות לנבדלותן בלקסיקון. בנוסף, נראה שאפילו הראיות מהמערכת האורתוגרפית של עברית, שבה חי"ת וכ"ף מיוצגות על ידי אותיות שונות, לא הספיקו על מנת לבסס את ההבחנה.

פרשנות אלטרנטיבית עשויה לגאול את הפונולוגיה האבסטרקטית בסגנון SPE. המילים שהוצגו בניסוי 3 היו כולן שמות עצם, שאין להם פרדיגמה ברורה, ולכן פחות מילים להשוואה בהן האלטרנציות השונות באות לידי ביטוי. במערכת הפועל, הפרדיגמה הרבה יותר נוכחת (למשל, בצורת העבר לעומת העתיד, כאשר כ"ף נמצאת בעמדה העיצורית הראשונה במילה, יש חילוף שיטתי בין [χ] ל[k]). מערכת הפועל מציגה הזדמנות למחקר המשך: ייתכן שבהקשר שבו החילוף נוכח בצורה יותר משמעותית, דוברים ממוזגים יהיו יותר מסויגים לגבי פריטים בהם כ"ף הוחלפה בחי"ת, בהסתמך על התפוצה השונה של העיצורים. תוצאה כזו תתמוך בייצוג מנטלי אבסטרקטי יותר של חי"ת וכ"ף במערכת הפועל.

## ממצאים הנוגעים לעיבוד שפה בזמן אמת

בנוסף להשלכות של ניסוי 3 על שאלות של ייצוג, התוצאות חשפו גם אפקט של עיבוד מבוסס-דובר. לניסוי הוקלטו שני דוברים: דובר ממזג, שלעולם לא הפיק חי"ת לועית; ודובר מבחין, שהפיק את שתי הקטגוריות, ובחלק מההפקות החליף חי"ת בכ"ף ולהפך. התוצאות עבור פריטים שהופקו בקול של הדובר הממזג היו דומות בקרב כל קבוצות המשתתפים.<sup>1</sup> לעומת זאת, מאזינים מרקע מבחין נטו גם לדחות או להגיב בצורה איטית משמעותית לפריטים בהם חי"ת הוחלפה בכ"ף, כאשר הדובר שהפיק את המילים נתפס ככזה שאמור להבחין בין הקטגוריות בהגייה. כלומר, פריט כמו maxog (מחוג בהגייה לא-מבחינה) היה מקובל על כלל המאזינים כאשר הופק בקול הממזג, אך נדחה על ידי המאזינים מרקע מבחין כאשר הופק בקול המבחין. הממצא הזה מדגים שהיכרות עמוקה עם דיאלקט יכולה ליצור ניבויים עדינים לגבי ההפקה הצפויה: אפילו כאשר פריטים היו דקדוקיים ברמה הלקסיקלית, כאשר הפקה לא היתה עקבית עם זהות הדובר היא נדחתה לעתים קרובות יותר ודרשה זמן עיבוד ממושך יותר. עבור מודל העיבוד הדו-מסלולי של קאי ושות' (Cai et al. 2017), ובו קלט לשוני מעובד במקביל במסלול לקסיקלי ובמסלול אינדקסיקלי (מבוסס-דובר), הממצא הנוכחי מספק עדות נוספת לטובת הטענה שעיבוד במסלול האינדקסיקלי מתבצע בזמן אמת.

פרק 4 עוסק באופן יותר ממוקד בסוג זה של עיבוד מבוסס-דובר בזמן אמת, בניסיון לברר עד כמה אוטומטי העיבוד במסלול האינדקסיקלי. הניסוי השתמש בגרסא אודיטורית של מבחן האסוציאציות החבויות (The Implicit Association Task; Greenwald et al. 1998). בפרדיגמה זו, משתתפת צריכה להגיב לארבעה סוגי גירויים משני סוגי קטגוריות שונים באמצעות אותם שני מקשים. כך למשל, בתחילת הניסוי יש להגיב לתמונות של פרחים ולתצוגה אורתוגרפית של מילים בעלות מטען חיובי (כמו "נהדר") במקש שמאל, ולתמונות של חרקים ומילים בעלות מטען שלילי באמצעות מקש ימין; בשלב מסוים, הקטגוריה של מילים בעלות מטען חיובי/שלילי מחליפה כפתור, כך שהמשתתפת צריכה עתה להגיב לתמונות של פרחים ולמילים בעלות מטען שלילי באמצעות אותו הכפתור. לאחר ביצוע המטלה בשני הקישורים השונים, מתקבל ציון שאמור לגלם את הקשר התפיסתי החבוי בין הקטגוריות, המבוסס על אחוזן הדיוק וזמני התגובה בבלוקים השונים (כלומר, כאשר פרחים קושרו למילים בעלות מטען שלילי לעומת חיובי). הציון המשוקלל הזה, שלוקח בחשבון גם את סדר ההצגה של האסוציאציות, נקרא d-score. במטא-אנליזה של 86 מחקרים שהשתמשו בפרדיגמה, נמצא שה-d-score הממוצע בקהילה בנוגע לקטגוריה מסוימת (למשל, אסוציאציה בין פרצופים של אנשים לבנים למילים חיוביות) עומד בקורלציה עם מגוון מדדים התנהגותיים (Greenwald et al. 2003).

מבחן האסוציאציות החבויות נבחר כיוון שהפרדיגמה כבר הוכיחה את עצמה כיעילה בחשיפה של תהליכי עיבוד לא מודעים. באמצעות העתקת המתודולוגיה לחומרים אודיטוריים בלבד, כיוונתי לבחון את האפשרות שעיבוד

<sup>1</sup> יוצא דופן אחד הוא משתתפים מקבוצת הדור הראשון למבחינים הגיבו לפריטים ממוזגים עם חי"ת אטימולוגית בצורה איטית יותר משתי הקבוצות האחרות. הממצא הזה מרמז על "מחיר" עיבודי למיזוג עבור דוברים לא-ממוזגים.

מבוסס-דובר גם הוא אוטומטי ולא מודע. המשתנים הפונולוגיים שנבחנו בניסוי 4 היו חיי"ת גרונית (לעומת הוריאנט הוילוני/ענבלי), והגיית הדיפתונג [ej] במילים כמו מימד (לעומת הגיית מונופתונג [e]), במסגרת שתי זהויות שונות שהוצגו למשתתפים כדובר בהקלטות: דובר מזרחי (שצפוי להשתמש בחיי"ת ובמונופתונג) ושדרן רדיו קלאסי (שצפוי להשתמש בחיי"ת ובדיפתונג). בניסוי 5, נבחנו המשתנים הפונולוגיים חיי"ת גרונית לעומת חיי"ת חיכת ביחס לרי"ש מתגלגלת ([r]). הזהויות שהוצגו למשתתפים בניסוי זה שתיהן של דוברי עברית כשפה שנייה: דובר ערבית פלסטינית (שצפוי להשתמש בחיי"ת גרונית וברי"ש מתגלגלת), ודובר ספרדית ארגנטינאית (שצפוי להשתמש בחיי"ת חיכת וברי"ש מתגלגלת). בכל הניסויים, המטלה היתה זהה: לקטלג את סוגי הגירויים, כאשר באמצע המטלה מקש התגובה מתחלף עבור אחד מהמשתתפים (למשל, אם בתחילה תגובה לחיי"ת ולדיפתונג היתה באותו המקש, בחלק השני התגובה לחיי"ת ולמונופתונג תהיה באמצעות אותו המקש).

בניסוי 4 נמצא קישור בינוני עד חזק בין חיי"ת גרונית להגייה בדיפתונג לפי מדד ה d-score, ללא קשר לזהות שהוצגה למשתתפים. בניסוי 5, נמצא קישור בין המשתנים בכיוון הצפוי רק כאשר הזהות שהוצגה למשתתפים היא "ארגנטינאית", אך לא כאשר הזהות המוצגת היתה "פלסטינית", וההבדל בין שני התנאים לא היה משמעותי. הדיון בפרק עוסק בהרחבה בסיבות האפשריות לכך שלא נצפה הבדל תלוי דובר באסוציאציות. ההמלצה העולה ממנו היא לחקור באמצעות מבחן האסוציאציות החבויות רק משתנים לשוניים מזוהים חברתית מספיק ומובחנים מספיק מבחינה אקוסטית, שיבטיחו גם זיהוי של הוריאנט עם הדובר, וגם קטגוריזציה פונטית טובה.

על אף שבמדד ה d-score לא נמצא אפקט של עיבוד מבוסס-דובר, ניתוח פוסט-הוק של אחוזי הדיוק וזמני התגובה חושף מגמה של אסוציאציה בכיוון הצפוי בניסוי 5 (חיי"ת חיכת+רי"ש מתגלגלת לדובר הספרדית, חיי"ת גרונית+רי"ש מתגלגלת לדובר הערבית). שני ממצאים רלוונטיים במיוחד למודל העיבוד הדו-מסלולי (Cai et al. 2017):

- א. קטגוריזציה של חיי"ת גרונית היתה מדויקת יותר תחת תנאי "הדובר הארגנטינאית", ביחס ל"דובר הפלסטינית". בשני המקרים, ההקלטות אותן שמעו משתתפים היו זהות, כלומר ההבדל נבע מתפיסת ההגאים ביחס לדובר.
- ב. זמני התגובה היו ארוכים יותר בתנאי שבו קושרו חיי"ת גרונית ורי"ש מתגלגלת תחת "הדובר הארגנטינאית". זהו הכיוון הצפוי של הקישור, שכן מבטא ספרדי-ארגנטינאי כולל את הוריאנטים חיי"ת חיכת ורי"ש מתגלגלת. ממצא זה נמצא בהלימה גם עם ה d-score בתנאי ה"ארגנטינאית".

ההסבר המוצע לממצא א' הוא שחיי"ת חיכת לא היתה מובחנת מספיק מבחינה אקוסטית עבור המאזינים. בהינתן שאין תווית אחת שמצביעה על הקטגוריה הזו בעברית, בניגוד לוריאנטים האחרים שנבחנו במטלה ("חיי"ת גרונית" ו"רי"ש מתגלגלת"), למאזינים היה קשה להתמקד בקטגוריה אליה כיוון הניסוי. תחת ה"דובר הארגנטינאית", הקטגוריה הפכה

לכולטת יותר; הידע המטא-לשוני של המשתתפים יכול היה לכוון אותם ("דוברים ארגנטינאים בדרך כלל מפקים חי"ת באופן מסוים, הנבדל אקוסטית מהאופן שבו חי"ת גרונית וכ"ף נהגות בעברית באופן הזה והזה") לחבר את הוריאנט החיכי עם התווית שבאמצעותה הוא יוצג בניסוי ("חי"ת רכה"). על ידי התמקדות במימד האקוסטי הרלוונטי, משתתפים תחת ה"דובר הארגנטינאי" הבחינו בצורה טובה יותר בין "חי"ת גרונית ו"חי"ת רכה", מה שהוביל לאחוזי דיוק גבוהים יותר בקטגוריזציה של חי"ת גרונית. ההסבר הזה נתמך על ידי השיפוטים הגלויים של משתתפים, על פי שאלון שמילאו לאחר המטלה: תחת ה"דובר הארגנטינאי", משתתפים נטו יותר לקשר את ההגייה החיכית עם דוברים ארגנטינאים. ממצא זה מדגים שידע מטא-לשוני יכול לעזור למקד את תשומת הלב של המאזינים למימד אקוסטי רלוונטי.

ממצא ב' מציע שזהות הדובר המוצגת השפיעה על מידת הקושי לקשר וריאנטים זה עם זה. מחקרים קודמים על עיבוד מבוסס-דובר הדגימו שהמגדר של דובר/ת והפרסונה שהיא משדרת יכולים להשפיע על גבולות בין קטגוריות פונולוגיות (Strand 1999; D'Onofrio 2018b). אך התוצאות האלה יכולות תמיד לקבל הסבר של כיוול מטא-לשוני מאוחר, כלומר, מאזינה יכולה לאמץ אסטרטגיה בסגנון "הצליל הזה נתפס אצלי כמו סוג של /a/, אבל הדוברת הוצגה בפניי כ"Valley Girl", אז עליי להיזיז מעט את הגבול - לכן אחליט שזו "æ". אפקט כיוול מטא-לשוני מסוג זה אינו יכול להיות המקור של התוצאות הנוכחיות: מכיוון שהמטלה זהה תחת שתי הזהויות שהוצגו למאזינים, ועל המשתתפות להגיב מהר ככל הניתן לכל גירוי, התוצאות משקפות אפקט כיוול בזמן אמת.

לסיכום, פרקים 3-4 תומכים במודל העיבוד הדו-מסלולי (Cai et al. 2017), לפיו מידע חברתי על הדובר יכול לתרום לעיבוד פונולוגי בזמן אמת. פרק 3 מדגים שהיכרות עמוקה עם דיאלקט יכול ליצור ניבויים עדינים לגבי הפקה, כמו לגבי השאלה האם דובר ישתמש בוריאנט ממוזג או לא. פרק 4 מוסיף שהצגה של דובר המשתמש בוריאנט מסוים יכולה למקד את תשומת הלב של המאזינים במימד האקוסטי הרלוונטי לקטגוריזציה, כמו גם להשפיע על הניבויים בנוגע לתפוצה של משתנים נוספים בהפקה של אותו הדובר.

אוניברסיטת תל-אביב  TEL AVIV UNIVERSITY

הפקולטה למדעי הרוח ע"ש לסטר וסאלי אנטין

ביה"ס לפילוסופיה, בלשנות ולימודי מדע



**לשמוע אתניות:**

**זיהוי, סטריאוטיפיזציה ועיבוד של מאפיינים פונטיים מסומנים חברתית בעברית  
מוזרנית**

חיבור זה הוגש לשם קבלת התואר "דוקטור לפילוסופיה"

באוניברסיטת תל אביב ובאקול נורמל סופרייר

על ידי

שיא ברבי

העבודה הוכנה בהדרכת

פרופ' אותי בת אל ופרופ' שרון פפרקמפ

אוקטובר 2021