2 Methodology

This study is innovative because it explicitly combines theoretic approaches which are applied on empirical data that was gathered in Israel. The main theoretic principles adhered to in this study belong to the areas of sociology, linguistics and cognitive science. Key notions which will be used in this study stem from the sociology of knowledge, as it is outlined in the works of Berger & Luckmann (1967) and Luhmann (1993), and perceptive variationist linguistics which is advocated for by Krefeld & Pustka (2010) and Purschke (2011). It will be argued from the perspective of cognitive linguistics that language is dependent on principles of categorization which have been discussed by Rosch (1978), Lakoff (1987), Langacker (2008) and Schmid (2020).

Although all these approaches are concerned with processes of categorization, they have – to the best of my knowledge – not been combined and reassessed in the light of an empirical study, so far. I argue that these theories can be used to build on each other to enrich all kinds of studies in the domain of sociolinguistics which has often been criticized as lacking a common theoretical foundation. Coupland (1998: 110) asserts against this criticism that all researchers in the field of sociolinguistics are "already theoretically engaged." However, I argue that the discipline would benefit from lying out and discussing its theoretical frameworks more explicitly in research reports, instead of assuming a more or less common theoretical basis. Due to the limited space for scientific articles and the focus on publishable results, it is not surprising that methodological discussions rarely find a place in research papers.

Another reason for the vagueness of the research area lies in its history and the term 'sociolinguistics' itself. The original American research paradigm was propagated in the form of usage-based, descriptive studies of linguistic variation, against the backdrop of the research tradition of theoretical linguistics or generative grammar (Sinner 2014: 11). Therefore, the stress on the social component in the denominator of the research area is understandable, from a historical perspective. Under these premises, sociolinguistics has come to harbor a multitude of subdisciplines, so diverse as language acquisition, second language learning, the study of language policy and language attitudes, linguistic forensics, variationist linguistics and dialectology, which by themselves include many subdisciplines.

This study is literally sociolinguistic because it investigates processes of categorization at the intersection of the social and the linguistic domain. At the same time, it is a study about human cognition and its social ramifications – the stronghold of cognitive linguistics – the second major paradigm in linguistics for the usage-based study of language. Since we view language as a social phenomenon which can only be studied based on its usage, it feels tautological to hint to this fact with terms like 'sociolinguistics' – any usage-based study of language has to deal with the

social dimension.[1] This definition excludes research which is strictly concerned with normative structures of a certain language, on a theoretical level – which Schulze describes as "meaningless noise."

> It is claimed that language is only language if it is perceived/processed by someone equipped with a linguistic knowledge system. Else, what we describe as language data is nothing but 'meaningless noise'. [...] language takes only place in terms of language practices. The reality of language practice is the sole source for obtaining language data as well as the most relevant factor for accessing linguistic knowledge. (Schulze 2012: 1–2)

In the following, the term 'usage event' will be used to designate instances of language use, in Schmid's (2020: 15) understanding of the term.

This study adheres to a non-objectivist outlook on language which is expressed concisely by Schulze:

> [L]anguage [can] be defined as a system of articulatory expression of cognitive states that is learned in its symbolization processes and socialized through imitation: It is a collective knowledge system that has been passed down. It is the expression of perceptions in experience, whose effectiveness is constructed as communication.[2] (Schulze 2010: 37, my translation)

Wherever it may be helpful for the argument, the social constitution of this knowledge will be reassessed with principles from the sociology of knowledge.

PVL, the main framework behind this study will be introduced in detail after a contextualization within dialectology and variationist linguistics. For the application of these theories in the research domain, the flexible and exploratory framework of grounded theory methodology which was originally presented in Glaser & Strauss (1967) was used in this study. I argue that empirical linguistic research can benefit from the systematic application of this research paradigm from the context of qualitative sociology.

It is not my aim to give an extensive overview about the research which has been carried out in the field of sociolinguistics and variationist linguistics, so far. An

---

1 Latour (2005: 5) criticizes the notion of 'social,' as in "social dimension," altogether and redefines sociology as the study of associations, which is in line with our understanding of sociolinguistic research: "Whereas sociologists (or socio-economists, socio-linguists, social psychologists, etc.) take social aggregates as the given that could shed some light on residual aspects of economics, linguistics, psychology, management, and so on, these other scholars, on the contrary, consider social aggregates as what should be explained by the specific associations provided by economics, linguistics, psychology, law, management, etc."

2 German original: Sprache [kann] definiert werden als ein in seinen Symbolisierungsverfahren gelerntes, über Imitation vergesellschaftetes – also kollektives, tradiertes und als Wissen gespeichertes, artikulatorisches Ausdrucksystem kognitiver Zustände – ergo Ausdruck von Wahrnehmungen in Erfahrung, dessen Wirksamkeit als Kommunikation konstruiert wird.

excellent overview is provided by Sinner (2014), Bokelmann (2020) and by Eckert (2012), with a focus on the American tradition. In the following, the most relevant theories and studies which served as a model for the present study will be reviewed.

2.1 Research on linguistic variation

The concept of 'linguistic variation' has been at the core of several research traditions within linguistics. Its theoretical basis goes back to structuralist linguistics and Saussure's concept of language as an abstract signifying system, with the linguistic sign as a symbolic pairing of form and meaning at its heart (Bokelmann 2020: 14). According to Bülow (2017: 36), Saussure's structuralist conception of language was already influenced by early system-theoretic thinking. The linguistic sign can only function within a system, in relation and in contrast to other signs (Bülow 2017: 38). Originally, structuralist linguists who saw it as their task to describe languages systematically were struggling with the fact that languages are not monolithic constructs, but subject to change – both from a diachronic and a synchronic perspective (cf. Bokelmann 2020: 21–22). Therefore, Saussure introduced the distinction of "langue," the system to be studied by the structuralists, and "parole," the actual language in its usage, which should not have a place in linguistics:

> The activity of the speaker should be studied in a number of disciplines which have no place in linguistics except through their relation to language. (Saussure 1959: 18)

Thus, all deviations from the ideal linguistic system (*langue*) in language usage (*parole*) could be treated as variation and practically ignored because it was defined as irrelevant to the description of *langue*. The concept of the 'allophone' served to integrate phonetic variability within structuralist theory: systematically differing phonetic realizations can still be grasped as a single phoneme.

In MH, /ʕ/ – represented by the letter *'ayn* – can be realized as [ʔ] and as pharyngealized [ʕ] without changing the lexical meaning of a lexeme that contains /ʕ/. From a structuralist perspective, [ʔ] and [ʕ] can be described as allophones of /ʕ/. This case of linguistic variation and its social meaning is the subject of Gafter's (2014) dissertation, which will be reviewed in the context of research on linguistic variation in MH in Israel (see 3.2.2).

What is commonly meant by 'linguistic variation' will be illustrated with another example. From a synchronic perspective, two speakers of the same language can express the same concept with different linguistic means. One case of variation in Hebrew is narrated in the Book of Judges:

> The Gileadites held the fords of the Jordan against the Ephraimites. And when any fugitive from Ephraim said, 'Let me cross,' the men of Gilead would ask him, 'Are you an Ephraimite?'; if he said 'No,' they would say to him, 'Then say shibboleth;' but he would say 'sibboleth,' not being able to pronounce it correctly. Thereupon they would seize him and slay him by the fords of the Jordan. Forty-two thousand Ephraimites fell at that time. (Judges 12, 5-6)[3]

One of the main interests of research on linguistic variation is about its meaning. This notorious instance of linguistic variation is described in the Bible as being socially conditioned: belonging to one social group, the Ephraimites, conditions the pronunciation of <ɯ> in *shibolet* 'ear of grain' as [s]. In variationist linguistics terminology *shibolet* and *sibolet* are 'variants' which belong to a 'linguistic variable.'[4] It is unknown if the Ephraimites pronounced /sh/ systematically as [s] or if *sibolet* was a single lexical variant. It is narrated that they could not say *shibolet* and that they could hence be identified as Ephraimites – even though they were denying their true identity. The Gileadites used a form of forensic linguistics.

There are areas of research on historic variables within linguistics. Historical linguistics, diachronic dialectology and diachronic variationist linguistics could investigate further into the historical variant *sibolet*. Today, this variable is no longer present in MH: the utterance of *sibolet* would probably point to an expressive language disorder of the speaker. Usually, there is just one possible realization, which is *shibolet*.

Besides meaning, another question drives research on linguistic variation, especially historical linguistics: how and why are certain variants propagated or disappear? Researchers try to uncover the dynamics and causes of linguistic change. Often, the reasons for diachronic linguistic change will remain buried, together with the speakers of the historic variants who can no longer be questioned, like the Ephraimites. Based on the context of the biblical narrative, one could argue that the dominant social group, the Gileadites, imposed their way of speaking by virtually eradicating all speakers of the variant *sibolet*.

Systematic phonetic variation – phonological variation – usually is a structural characteristic of 'accents' or 'dialects.' Had the Ephraimites consequently pronounced every <ɯ> as [s], it would seem sensible to speak of an Ephraimite 'accent.' If they had consequently varied in their language production from neighboring people in a considerable number of linguistic variables on the phonetic, the lexical and even the morphosyntactic level, one would be inclined to speak of an Ephraimite 'dialect.' Instead of 'dialect,' one could have used the notions of 'sociolect' or 'ethnolect' if their

---

3 Quoted from https://www.sefaria.org/
4 According to Chambers & Trudgill (1998: 13) *shibbolet* is a loan word in several languages and means "test word" or "a distinguishing trait." Thus, it stands metonymically for indexical variants.

activities were not restricted to a certain geographic region because of a nomadic life style, for example. Had the Ephraimites managed to acquire considerable political and cultural power – especially by the means of sustainable literary products, one would certainly have heard of an 'Ancient Ephraimite language.'

As Kristiansen points out, the integration of linguistic variation into a theoretical linguistic framework is challenging and contradictory approaches are co-existing:

> It might be argued that what is at work are two different levels of granularity: socially indexical phonetic variants [such as the Ephraimites' *sibolet*] pertain to the level of lectal varieties (and hence to dialectology), phonetic variation in general to phonology – but that type of clear-cut division into structural levels is not in consonance with the basic tenets of Cognitive Linguistics, which prefers to consider meaningfulness in terms of constructions and assemblies across structural levels. (Kristiansen 2008: 72)

The approaches to linguistic variation which are most relevant for this study and their outlook on the meaning of variation, as well as their key concepts will be reviewed in the next sections.

2.1.1 Dialectology

As a European linguist I am most familiar with dialectologist approaches to linguistic variation, referred to as "dialect geography" by Chambers & Trudgill (1998: 13), which aim to represent linguistic data in conjunction with geographic data in the form of maps or linguistic atlases.

Early dialectology developed in 19th century Germany under the influence of the so-called neogrammarians (Chambers & Trudgill 1998: 14). A short review of their research program is due because they were the first linguists to leave the office and to study language based on its usage. The original aim of the early German dialectologists was to gather empirical evidence for their theory of *Lautwandel* 'sound change.' With works such as *Prinzipien der Sprachgeschichte* (Paul 2010 [1880]) they were striving not just to describe, but to explain *Sprachentwicklung* 'the development of language' (Jungen & Lohnstein 2007: 179). The neogrammarians' dialectology was inspired by an estimation for the vernacular language which had not received scholarly attention (Jungen & Lohnstein 2007: 177–178). It is likely that their positive outlook on the vernacular was inspired by romanticist ideas of their time – such as the quest for truth in nature and accordingly for the true language among the people. It was also at the beginning of the 19th century when scholars set out as wanderers and collected popular tales (*Volksmärchen*) from elderly rural speakers. These tales, like Runge's *Van den Fischer und siine Fru* contained vernacular traits in

their edition by Grimm & Grimm (1812).[5] However, the neogrammarians advocated a strictly scientific position, against all metaphysicist and biologist theories of their contemporaries (Jungen & Lohnstein 2007: 179–180). Remarkably, they resorted to the young discipline of psychology to explain inconsistencies in their theories about sound change (Jungen & Lohnstein 2007: 180). Chronologically, the short-lived neogrammarian tradition was followed by structuralism and with Ferdinand de Saussure, who became the iconic structuralist, the focus of linguistic research shifted back from *parole* to *langue*, for at least half a century. Saussure had studied in Leipzig under the neogrammarians Leskien, Osthoff and Brugmann (Jungen & Lohnstein 2007: 184).

The methods which were used in dialectologist projects developed in the course of a century, but essentially all adhere to the same concept: in their work with participants, dialectologists have been aiming to elicit 'natural speech.' The choice of participants was determined by the assumption that "nonmobile, older, rural male" (NORM) participants (Chambers & Trudgill 1998: 29) would "use the most conservative variety of a language" (Schneider & Barron 2008: 16). This focus on the prototyical NORM participant was criticized by Chambers & Trudgill (1998: 29) as well as by others (Bisang 2008: 15 and Schneider & Barron 2008: 16): it has been used in many studies irrespective of varying circumstantial factors, such as "culture" and "socioeconomic climate." For the data collection which can take place as interviews or written surveys, questionnaires have been used to allow for comparable data (Chambers & Trudgill 1998: 21–22). It has been a huge challenge to analyze the wealth of data and to represent it concisely (Chambers & Trudgill 1998: 16). Often, they have been focusing on the phonetic domain because it was best observable. Based on the earliest maps by Georg Wenker, which date from 1881, many linguistic atlases were published (Chambers & Trudgill 1998: 16–20).

To the best of my knowledge, a similar research project has never been completed for Hebrew in Israel. At the same time, there is a considerable research tradition on Arabic dialects in Israel and the West Bank with at least one recently published linguistic atlas (see Behnstedt & Kleinberger 2019). In contrast, linguistic research on MH has been preoccupied primarily with *langue*, while usage-based studies have been somewhat neglected. The reasons for this theoretical orientation are partly ideological and will be discussed together with recent empirical studies which have been shifting the focus to usage in 3.2.2.

---

5 The introduction contains this citation to convey their conviction: "Man sollte die Weisheit der Völker, bei denen man lebt, in ihrer mannichfaltigen Gestalt, selbst in Liedern, quas ad ignem aniculae narrant puellis, aufspüren und in Umlauf bringen. (Histor. Critik I. 245.)" (Grimm & Grimm 1812: XXIII)

Finding myself in a pioneer position at the beginning of my research, it did not seem reasonable to stick to the methods which have been used in traditional dialectology, for two reasons: firstly, a dissertation is not the right format to undertake a large scale dialectologist project in an under researched area because it does not allow for sufficient financial, timely and human resources. Secondly – and more importantly – Israel does not seem to qualify as a promising field for dialectologist research (except for research on Arabic) because of the particular history of MH which is seen by many as a 'revived language.' As lined out above (see 1.1), I was convinced that there have to be more crucial variables than the mere changing of a geographic position, when it comes to determine the nature of linguistic variation in MH in Israel. A focus on participants with the NORM characteristics would not have served to capture the fine grained differences in the use of MH I had in mind.

2.1.2 Studies on sociolinguistic variation

As outlined above, sociolinguistics comprises many linguistic subdisciplines, today. It originated as the investigation of sociolinguistic variation – more precisely – of socially marked variants. From a bulk of research, which is subsumed by Eckert (2012: 3) under the term "first wave of variation studies," it is well known that the interplay of the regional and the social dimension can have linguistic effects. In the introduction of his paper, entitled *Phonological Correlates of Social Stratification*, Labov positions his research program in contrast to the structuralist tradition:

> AS WE approach the study of language in its social context, it seems that by the very same steps we enter the study of small differences in language behavior. For many years, the structural analysis of sound systems has enjoyed, and profited by, a kind of bold abstraction from such differences. Small differences within a system have been explained away as "free variation" or "social variants," and we have concentrated on the abstract organization of constant features. But to understand the dynamics of such systems, the mechanism of their evolution, and their role in community life, it is useful to reverse this attitude. (Labov 1964: 164)

Labov's motivation can be compared with the neogrammarians' aim: to study language based on its usage and thereby gather empirical evidence for the description and the explanation of linguistic change. The methods used for the collection of data are also based on the "dialectological tradition" (Labov 1984: 28). What is remarkable about studies on sociolinguistic variation is the subsequent narrowing of the research focus from the dialectologists' domain of 'regionality' to notions of 'social classes,' 'gender,' 'age,' 'ethnicity' and to complex constructs such as 'identities' and 'styles.' Therefore, the researched population is treated as 'social groups' or 'net-

works' in approaches which are informed by ethnography and qualitative sociology. Kristiansen summarizes Labov's approach as follows:

> Labov's model, often referred to as the 'attention paid to speech' model, investigates how speech styles vary according to situation. Assuming that speakers adjust their speech to contextual factors whenever they are aware of these, attention can be viewed as a determining factor which links linguistic variants to social variables. Labov aimed at eliciting 'real' unmonitored speech by drawing the speaker's attention away from situational factors. (Kristiansen 2008: 73)

With the term 'observer's paradox,' Labov expressed his sensitivity towards the artificiality of situations when one's speech is observed for linguistic analysis (Figueroa 1994: 90–91). One of his key methods is the 'sociolinguistic interview' which is aimed to yield recordings of up to two hours of speech by eliciting "narratives of personal experience, where community norms and styles of personal interaction are most plainly revealed, and where style is regularly shifted towards the vernacular" (Labov 1984: 32).

2.1.2.1 Variants as markers of macro-sociological categories

According to Eckert (2012: 2), first wave studies were large scale surveys which "laid a solid foundation for the study of variation by establishing broad correlations between linguistic variables and the macro-sociological categories of socioeconomic class, sex class, ethnicity and age." The participants' age was considered to be a promising factor for the study of linguistic change, based on the conviction that speakers tend to conserve their own linguistic system over time (Eckert 2012: 5).

In his New York City study, Labov (2006 [1966]: 41) argues for a correlation between variation in the post vocalic realization of /r/ and the social stratification of New York City speakers. For the study, salespeople in three socially distinct New York department stores were surveyed. The study's hypothesis is based on the notion of the "socio-economic class" (Labov 2006 [1966]: 130). Milroy & Milroy (1996: 52) and Coupland (1998: 114) assert that Labov draws on the work of Parsons (1991 [1951]) and his concept of "stratificational social class." Parsons also uses the concept of 'prestige' for his theory of social stratification:

> This ranking system in terms of esteem is what we may call the system of stratification of the society. [...] many elements of the "style of life" come to have significance, among other things, as symbols of prestige in the system of stratification. (Parsons 1991 [1951]: 89)

In first wave studies, 'prestige' has been regarded as a quality of certain variants – linguistic symbols of prestige – because they are associated with prestigious speakers (Eckert 2012: 3). Schmid defines 'prestige' as follows:

> Both in general and with regard to language, prestige is a positive inter-individual social
> evaluation. Prestige is connected with influence; lack of prestige is associated with deference
> and can therefore become stigmatized. Importantly, the behaviour of those who have status
> and prestige is likely to be imitated by those who do not and especially by social aspirers
> located somewhere in the middle of social hierarchies and looking upwards for role models. [...]
> Speakers are more likely to adopt utterance types and their variants from prestigious groups
> and speakers than from those who lack prestige and are stigmatized. (Schmid 2020: 114)

Labov documented the imitation of prestigious speakers in his New York City study:
the surveyed salespeople adapted their realization of /r/ to the realization of pres-
tigious customers. Thereby, he refuted the relativistic Bernstein hypothesis which
argues for the social determination of language use by the over-simplistic juxtaposi-
tion of low social status with a restricted code (Sinner 2014: 12). With the description
of the linguistic variable /r/, in post vocalic position, as being closely tied to the
"socio-economic differentiation" of the society, Labov (2006 [1966]: 41) suggests a
somehow conventionalized view on hierarchic relations within this differentiation
(Eckert 2012: 3). In other words, a meaningful linguistic imitation of prestigious
speakers is dependent on a common awareness of the socio-economic differences.
Moreover, speakers need to agree on a hierarchy which relates these differences
to the notion of 'prestige' and they need to agree on the symbolic relation of this
hierarchy with patterns of linguistic variation. Since notions of 'prestige' seem to
be conventionalized as related to patterns of linguistic variation, these notions can
be explored to gain insights about linguistic variation. This line of reasoning is the
basis for the elicitation task GERT which was used in this study (see 4.2.6).

Eckert points out that Labov's estimation of the vernacular was politically moti-
vated:

> Sociolinguists will agree that the political economy is fundamental to variation, and it is im-
> portant to recall that Labov's focus on social class was a significant political-academic move
> (as was his later focus on race). But the abstracted socioeconomic hierarchy provides only a
> general roadmap to the sites of linguistic production. (Eckert 2012: 6)

Consequently, the sociolinguistic researchers' interest in 'social class' as a macro-
social variable causing linguistic variation was diverted to other domains. In the
second wave studies, the focus was further narrowed on particular groups of speak-
ers who were studied in detail, with ethnographic methods, to get a more holistic
understanding of the causes for variation. Eckert (2012: 7) classifies the ethnographic
study of a community of fishermen on the island Martha's Vineyard (Labov 1963) as
a second wave study – although it preceded the New York City study. Second wave
studies, such as the seminal study of social networks in Belfast by Milroy (1980),

are characterized by Eckert (2012: 11) as focusing on "configurations rather than categories."

Both first and second wave studies treated variants as markers of (local) identity which are directly connected to their users who were treated as social groups or networks (Sinner 2014: 15). The notion 'social group' has been criticized as potentially vague (Sinner 2014: 143). Neuland & Schlobinski (2017: IX) argue that the notion is most often used implicitly in sociolinguistics, without defining its scope. Moreover, it would be misleading to conceptualize social groups and varieties as static entities (Neuland & Schlobinski 2017: XI). However, the advantage of this notion is its universal use which is not restricted to linguistics: it designates a social entity which ranges in size between individuals and nations. While 'nation' is a very abstract concept, 'social group' is closer to the participants' *Lebenswelt* and can therefore be used as a stimulus for elicitation, in tasks or interviews. Despite its conceptual vagueness, the notion can be used for investigations into the speaker's own systems of categorization – instead of applying a priori categories, such as 'social class.'

Inherent in all of these studies of sociolinguistic variation is the theoretical distinction between 'standard' and 'non-standard' variants.

> The term standard has been used to refer to speech that lacks clear regional and/or socially stigmatized features – the variety legitimized by, and required for meaningful participation in, institutions of education and economic and political power. This is the variety typical of the educated upper middle class. The assumption from the start has been that language varieties carry the social status of their speakers, making the class stratification of language a continuum of linguistic prestige. (Eckert 2012: 3)

Although studies on sociolinguistic variation could investigate only a few variables, in certain contexts and moments, their underlying goal was to find out more about how these variables relate to each other in dynamic linguistic systems. The conception that several 'language varieties' – at least, more than one – are available in every speech community is a premise of the research paradigm. Theories about the organization of the linguistic system, based on varietal subsystems, and the classification of varieties are the subject of 'Varietätenlinguistik' (see 2.1.3).

2.1.2.2 Language and the social construction of meaning

First and second wave studies on sociolinguistic variation backed the hypothesis that language can have a constitutive function for social groups: it can be used to construct groups and to demonstrate belonging (Sinner 2014: 147). This is the basis for the third wave of variation studies which asks about the meaning of certain variants for their speakers. Eckert (2012: 14) describes third wave studies as the products of a shift from the "study of structure to the study of practice." In comparison to

the first and second wave studies, the focus is further narrowed on the individual speaker who is now attributed with agency. The speakers' use of linguistic variation to construct personal and social styles – 'roles' in Berger & Luckmann's (1967) terms (see 2.1.2.3) – is at the heart of third wave studies.

Several basic assumptions, which are more or less tacitly expressed in these studies, can be traced back to sociological research paradigms, such as the sociology of knowledge and qualitative sociology (cf. Bülow 2017: 87). These paradigms, which are attributed to constructivism, gained momentum in the second half of the twentieth century, foremost in the USA. Charmaz, who contributed to the development of GTM, elaborates on the understanding of 'meaning,' in the research paradigm of symbolic interactionism:

> People confer meanings on things – whether these things are objects, events or people. Meanings do not inhere in things as individuals ordinarily assume. Nor are meanings singular and shared by all. Instead, meanings are multiple and situated in specific contexts. What you do with something arises from what it means to you – and these meanings have consequences. (Charmaz 2004: 58)

The "things" that are investigated by third wave studies are primarily linguistic variants and the categories which the speakers associate with them. According to constructivist principles, it is necessary to inquire into the speakers' own systems of categorization for any analysis of meaning. Therefore, it is central for third wave studies to yield insights into the interdependent processes of categorization of social groups and of linguistic variation – a research goal of this study which is framed by RQ8 and RQ9.

Context sensitivity and the subjectivity of 'meaning' are a methodological challenge. While the methods of third wave studies are essentially the same which have been used in the second wave (ethnographic fieldwork and interviews), data analysis is guided by principles which stem from qualitative sociology. Among third wave studies, there is no consensus about the application of these principles. Different frameworks are used and combined – at times without a transparent methodological basis. Besides GTM, other approaches have been systematized in qualitative sociology, such as the highly structured qualitative content analysis by Mayring (2015) and Kuckartz (2016).

The following is no attempt to explain the function of language in the social construction of meaning – several monographs are devoted to this complex topic (e.g. Eckert 2018 and Harder 2010). Instead, these processes will be reassessed in the light of three approaches, which are most relevant for this study: Berger & Luckmann (1967) is a seminal work in the sociology of knowledge which has been ground-laying for social constructivism, Schmid (2020) is an encompassing theory about *The Dynamics of the Linguistic System* which is based on cognitive linguistics within a

system-theoretic framework and Levon (2010) will be reviewed as an example for a third wave study, in 2.1.2.4.

Berger & Luckmann define the research area of the sociology of knowledge by posing the "central question for sociological theory:"

> How is it possible that subjective meanings become objective facticities? [...] an adequate understanding of the 'reality sui generis' of society requires an inquiry into the manner in which this reality is constructed. This inquiry, we maintain, is the task of the sociology of knowledge. (Berger & Luckmann 1967: 30)

When transferring this agenda to the research on linguistic variation, the "objective facticities" which have to be analyzed refer to linguistic variants, like the Ephraimites' *sibbolet*, and the variationist's task becomes the inquiry into the manner in which the semantic relations between variants and concepts, such as 'Ephraimites,' are constructed. Conventionalized associations with variants are referred to as "metapragmatic stereotypes" by Agha (2006: 148) who analyses these conventionalization processes as "enregisterment."

The central premise of the sociology of knowledge is the treatment of "objective facticities" as social constructions. This may look like a contradiction in itself because it does not fit the common understanding of objectivity as completely independent of the social domain. Nevertheless, it resonates with our "non-objectivist" understanding of language as a socially created system of knowledge (Schulze 2012: 1; see above 2). In fact, Berger & Luckmann's "central question for sociological theory" touches upon the groundwork of linguistic theory: to get an adequate understanding of language, the processes of construction that are underlying Saussure's linguistic sign have to be investigated. Third wave studies are driven by a similar question: How is it possible that subjective meanings become linguistic signs?

Leaping one step ahead of this fundamental question, Berger & Luckmann emphasize the important function of language in the construction of reality:

> The language used in everyday life continuously provides me with the necessary objectifications and posits the order within which these make sense and within which everyday life has meaning for me.[...] In this manner language marks the co-ordinates of my life in society and fills that life with meaningful objects. (Berger & Luckmann 1967: 22)

Furthermore, they regard language not only as crucial for conveying meaning in social interaction – e.g. to demonstrate belonging to social groups – but also for constructing one's own identity.

> It can [...] be said, that language makes 'more real' my subjectivity not only to my conversation partner but also to myself. [...] This very important characteristic of language is well caught

in the saying that men must talk about themselves until they know themselves. (Berger & Luckmann 1967: 38)

The above statement is reminiscent of Goffman's (1956) conception of social inter-action as acts of performance to present oneself. This treatment of language as a means of performance is inherent in third wave studies, which revolve around the notion of 'identity' – whereas the first wave studies were primarily concerned with the notion of 'social class'. Schmid even affirms the importance of 'identity' as a force within the linguistic system, from a general perspective:

Identity is an important force acting directly on usage and indirectly on the social and cognitive processes that lead to the sedimentation of linguistic structures. (Schmid 2020: 40)

The basic difference of first and second wave studies, in contrast to the third wave, is that variants are no longer understood as subconscious and involuntary expressions of a speaker's social identity (i.e. as markers of identity) which is determined by several factors, such as membership in a particular social group (Sinner 2014: 15 and Buchholtz & Hall 2004: 382). Resorting to the terms of Berger & Luckmann, linguistic variants have to be studied in conjunction with their underlying "subjective meaning" and not just as "objective facticities" which are determined by objective extra-linguistic variables such as 'social class,' 'age' and 'sex.'

These extra-linguistic variables which were looked upon as independent from the speakers' agency can no longer be treated as causal triggers for linguistic varia-tion. Instead, the categories have to be investigated that are used by the speakers themselves, in processes of meaning construction. It is assumed that the function of linguistic variation can only be understood in relation to these processes. Thus, introducing speakers' agency into the variationist framework brings on an updated concept of linguistic variation, its meaning and its usage.

[L]anguage should be understood as providing a pool of resources – semiotic tools that indi-viduals can make use of in the variable performance of identities. Particular ways of talking are not essential components of individuals, but are, instead, social/ideological artifacts that people can take up and put down as the need or desire arises. Yet these sociolinguistic tools are not always ready for the taking. Rather their socially licit use tends to be strictly policed by dominant norms of a given society. The central focus of research in this paradigm, then, is to examine how individuals negotiate these ideological imperatives, using the social resources available to them to create the identity performances they desire. (Levon 2010: 65)

In this line of argumentation, Schmid elaborates further on the meaning of "ut-terance types," which are understood as variants in our context (see 2.1.4.1 for a terminological discussion):

we should avoid thinking about conventionalized utterance types as ready-made, off-the-peg tools to be grabbed and exploited in a customary manner for reaching equally unshakable communicative aims. Instead, they should be regarded as co-semiotic potentialities competing for dominance in a multidimensional probability space. [...] It lies in the very nature of utterance types as parts of this abstract system that they do not occupy a specific and fixed location in the probability space. Instead, they keep reconfiguring the terrain they can cover in a way which is at the same time extremely flexible and constrained by their usage history. The processes that distil conventions from usage histories and keep adapting them are usualization and diffusion. (Schmid 2020: 92)

While Levon focuses on normative aspects as acting forces during the processes of negotiation, Schmid's focus lies on the "dynamics of the linguistic system" and the functionality of its processes, which he frames as "usualization" and "diffusion." For a better understanding of these processes, they will be reassessed from a sociological perspective. Therefore, Luhmann's (1993) contextualization of the sociology of knowledge within a system theoretic framework is helpful because it brings together the notions of language as a system of knowledge and its social constitution. Luhmann presupposes the existence of alternative possibilities as a premise for all meaningful social action:

Our starting point is that all human experience and action takes place in a meaningful way and is itself only accessible through meaning. This means that the object of the intention and realization of the current execution is only given in the form of a reference to other possibilities.[6] (Luhmann 1993: 17, my translation)

This premise is taken up by Schmid (2020: 92) as a "multidimensional probability space." In theories of linguistic variation, this probability space is understood as composition of linguistic variables and their possible realizations as variants. The system-theoretic conception of language relates the realization of linguistic variants to all other possible realizations within a linguistic variable. Luhmann elaborates further on the conditions of purposeful interaction:

Every meaning contains a kind of guarantee of connectivity for further experience and action and a guarantee of recurrence, of a return to itself after passing through other meanings. Meaning therefore presents what is real interspersed with other possibilities and thus places behavior under pressure of selection because only one or the other eventuality of this pre-

---

6 German original: Unser Ausgangspunkt ist, daß alles menschliche Erleben und Handeln sinnförmig abläuft und sich selbst nur sinnförmig zugänglich ist. Das heißt, daß das, was jeweils Gegenstand der Intention und Realisationskern des aktuellen Vollzugs ist, nur in der Form der Verweisung auf andere Möglichkeiten gegeben ist.

sented surplus of possibilities can be currently realized, thematically intended, action-wise comprehended.[7] (Luhmann 1993: 17–18, my translation)

In a usage event, the speaker who has access to a pool of linguistic variants has to select exactly one variant for each variable. In addition to this structurally determined *Selektionsdruck*, social pressure weighs on the speaker's choice – as Levon (2010: 65) pointed out in reference to the "dominant norms of a given society." Luhmann (1993: 48) defines the process of selection as determined by additional factors which "transform the mere appearance of variation into a semantic career."[8] Hence, the speakers' agency is partly determined by interpersonal processes, which result in 'linguistic norms' through processes of "usualization" (Schmid 2020: 5, 40, 98).

2.1.2.3 Institutionalization and linguistic norms
Schmid's definition of usualization as a process that produces linguistic norms, is based on Berger & Luckmann's (1967) notion of 'institutionalization' (Schmid 2020: 111). Bülow (2017: 50) points out that Wittgenstein already compared linguistic rules to men made institutions, in his *Philosophische Untersuchungen*. Saussure (1916: 33) also used this analogy, in his statement: "la langue est une institution sociale." However, his understanding of 'institution' is more static than the one adhered to by the frameworks which are reviewed here.

Berger & Luckmann (1967: 70) presuppose that "all human activity is subject to habitualization." With a focus on its function, Berger & Luckmann (1967: 71) describe the process of "habitualization" as a "psychological gain" because it implies the narrowing of choices for the benefit of cognitive economy. Luhmann similarly defines the process of "Stabilisation" as the systematization of knowledge which enables its passing on through learning and can lead to semantic tradition:

> In the area of semantic tradition, this function is fulfilled by systematizing and dogmatizing knowledge. The resulting forms and types are reduced to more abstract rules that can be learned and passed on in a simplified form and can organize larger numbers of cases. Institutio

---

7 German original: Jeder Sinn enthält damit eine Art Anschließbarkeitsgarantie für weiteres Erleben und Handeln und eine Garantie für Rekurrenz, für Rückkehr zu ihm selbst nach Durchlaufen anderer Sinngehalte. Aller Sinn [18] präsentiert deshalb Wirkliches durchsetzt mit anderen Möglichkeiten und setzt das Verhalten damit unter Selektionsdruck, weil von diesem appräsentierten Möglichkeitsüberschuß nur die eine oder die andere Eventualität aktuell realisiert, thematisch intendiert, handlungsmäßig nachvollzogen werden kann.

8 My translation from the German original.

is a Roman term: It designates the connection between semantic order and doctrine.[9] (Luhmann 1993: 50, my translation)

Berger & Luckmann describe the individual process of habitualization as the basis for institutionalization:

> Institutionalization occurs whenever there is a reciprocal typification of habitualized actions by types of actors. Put differently, any such typification is an institution. [...] The typifications of habitualized actions that constitute institutions are always shared ones. [...] The institution posits that actions of type X will be preformed by actors of type X. (Berger & Luckmann 1967: 72)

In our context, knowledge about linguistic variation can be understood as an institution which posits that people of type X speak in manner X – or, more precisely, with the variants X, Y and Z.

Referring to our biblical example, there must have existed an institution among the Gileadites which posited, among other things, that *sibolet* is a typical variant that will be uttered by actors of the type 'Ephraimite.' Actually, it is not known if the terms 'Gileadites' and 'Ephraimites' were used by the groups themselves (as endonyms), or if they were used by others (as exonyms) and possibly introduced in retrospective for narrative purposes. Therefore, the institution among the Gileadites might have posited the basic distinction between 'us' and 'them' and the relation between 'them' and *sibolet*.

It is important to point out that an institution is independent of the truth of its typified relations between actions and actors. That is to say, despite the existence of the described institution among the Gileadites, the majority of the Ephraimites might have used the variant *shibolet*. Or the speakers who actually used *sibolet* might not have self-identified as Ephraimites. The relation between linguistic knowledge and linguistic production will be elaborated in 2.1.4.1.

Berger & Luckmann attribute language with an institutionalizing function which serves the speakers to create order:

> [L]anguage objectifies the world, transforming the *panta rhei* of experience into a cohesive order. In the establishment of this order language *realizes* a world, in the double sense of apprehending and producing it. Conversation is the actualizing of this realizing efficacy of language in the face-to-face situations of individual existence. (Berger & Luckmann 1967: 173)

---

9 German original: Im Bereich semantischer Tradition wird diese Funktion durch Systematisierung und Dogmatisierung des Wissens erfüllt. Die anfallenden Formen und Typen werden auf abstraktere Regeln gebracht, die vereinfacht gelernt und tradiert werden und größere Fallmengen ordnen können. Institutio ist der dafür zuständige römische Begriff. Er bezeichnet den Zusammenhang von semantischer Ordnung und Lehre.

Again, the Gileadites made use of the concept 'Ephraimites' to order their social environment into 'us' and 'them.' It is claimed that the ordering function of language is only effective through endless processes of 'conversation' because by its nature, the created order is futile. In conversation, institutions are acted out and experienced through the performance of 'roles' (as described by Goffman 1956). Berger & Luckmann define 'roles' as "types of actors" which are prior to institutionalization – like 'Ephraimite,' in our example:

> By playing roles, the individual participates in a social world. By internalizing these roles, the same world becomes subjectively real to him. In the common stock of knowledge there are standards of role performance that are accessible to all members of a society, or at least to those who are potential performers of the roles in question. (Berger & Luckmann 1967: 91)

These "standards of role performance" enable the judgment of performances as compliant or non-compliant with role standards (Berger & Luckmann 1967: 92). Thus, a Gileadite who uttered the non-standard variant *sibolet* might have been judged as non-compliant with his role and even be subjected to punishment by his fellows.

Taken back to the linguistic perspective on conversation, a certain degree of semantic stability can be reached by the repetition of patterns – 'usualization' in Schmid's (2020) terms. From a holistic perspective on the linguistic system, linguistic norms are expressions of the speakers' preference of certain variants over others, in certain situations, and can indicate the probability of their selection. In variationist theory, linguistic norms are understood as marking the distinction between standard and non-standard variants. The implicit norms which shape the representations of this standard among the speakers are commonly explicated through further processes of negotiation which revolve around the notion of 'correctness' of language (Schmid 2020: 98). Concepts such as *Hochdeutsch* (as 'normative correct German') and *'ivrit tiḳnit*, 'correct Hebrew,' are products of these negotiations.

Bourdieu highlights the role of social power when describing these negotiations in pedagogic institutions:

> It follows that the legitimate language is a semi-artificial language which has to be sustained by a permanent effort of correction, a task which falls both to institutions specially designed for this purpose and to individual speakers. (Bourdieu 1991: 60)

According to Bourdieu (1991: 61) it is paradoxical that pedagogic institutions are teaching rules for the practical act of speaking which grammarians have extracted retrospectively from the writings of professionals for their works of the codification of the language. Because of the artificial nature of this process of codification, knowledge of these normative rules can only be acquired in specific contexts. In this line

of thought, the institutionalized confinement of this knowledge to a certain group of people is one reason for its association with prestige:

> [T]he correct, i.e. corrected, expression owes the essential part of its social properties to the fact that it can be produced only by speakers possessing practical mastery of scholarly rules, explicitly constituted by a process of codification and expressly inculcated through pedagogic work. (Bourdieu 1991: 61)

In summary, Berger & Luckmann describe the relation between roles and institutions:

> Looked at from the perspective of the institutional order, the roles appear as institutional representations and mediations of the institutionally objectivated aggregates of knowledge. Looked at from the perspective of the several roles, each role carries with it a socially defined appendage of knowledge. [...] society exists only as individuals are conscious of it [...] the institutional order is real only in so far as it is realized in performed roles and that, on the other hand, roles are representative of an institutional order that defines their character (including their appendages of knowledge) and from which they derive their objective sense. (Berger & Luckmann 1967: 96)

As Schmid points out:

> Not only institutionalized social roles but also socially constructed categories such as gender, ethnicity, or age are established, perpetuated, and modified in this way. (Schmid 2020: 112)

Importantly, this has implications for the treatment of all of these categories in third wave studies. The analysis of any institutional is essentially an analysis of the members' knowledge (Berger & Luckmann 1967: 82). Based on these observations, 'meaning' is understood as constructed by speakers in specific contexts in relation with 'roles' which are representing 'institutions.' 'Identity' is understood as manifestations of meaning, as it is bestowed by individuals upon their own life and their relations to others which are apprehended by the performance of roles and in relation to institutions. Typified identities can in turn be apprehended as roles. Our working definition resonates with this somewhat technical definition:

> *[I]dentity*: an outcome of cultural semiotics that is accomplished through the production of contextually relevant sociopolitical relations of similarity and difference, authenticity and inauthenticity, and legitimacy and illegitimacy [...] (Buchholtz & Hall 2004: 382)

In general, there is an abundance of forces which can affect institutionalization (Luhmann 1993: 48). With a focus on the linguistic system, Schmid (2020: 111) presents a systematic overview of the "forces affecting conventionalization" of utterance types. The multitude of these factors and the complexity of these concepts makes their investigation a huge methodological challenge (Kehrein et al. 2010: 377–378).

Having laid out the theoretic foundation for third wave studies on linguistic variation, one exemplary study shall serve to illustrate how these complex concepts can be explored.

2.1.2.4 Linguistic variants and categories of sexuality in Israel

Levon's (2010) study about language use in lesbian and gay communities in Israel presupposes that meaning is "a property of both already existing social structures and individual constructions of identity" and consequently, that its construction is a process of negotiation (Levon 2010: 9). Levon (2010: 164) argues against the sole relevancy of the factor 'sexuality' for the description of "sexual subjectivities of Israeli gays and lesbians, or the linguistic practices through which those subjectivities can be socially realized." He does so by analyzing the use of morphological gender in plural forms, among other structures. His analysis is based on a corpus of spoken Hebrew which he compiled with recordings of his own sociolinguistic interviews. He treats the Hebrew masculine and feminine plural pronouns *hem* and *hen* as variants and argues for the masculine pronoun as the standard form for reference to multiple people, regardless of their sex (Levon 2010: 145).

In normative Hebrew grammar, reference to groups of women requires the feminine pronoun *hen* and the feminine plural morpheme *-ot*, whereas the presence of one man in the group would trigger the use of the masculine pronoun *hem* and the corresponding morpheme *-im*. For example, a group of teachers can be referred to by *morot* 'teachers.F' and *morim* 'teachers.M' (cf. French *enseignants* vs. *enseignantes*).

In Levon's sample of self-identifying homosexual participants, the women who affiliated with radical political activist groups were more likely to use masculine morphology, compared to the women who did not.

> [T]he women on the whole use a greater proportion of gender-specific morphology that the men do, and frequently in interviews I observed the women consciously self-correct from a generic masculine form to a (feminine) gender specific one. I would argue then that the systematic use of masculine morphology among the Radical women when speaking about out-group referents is likely something more than a distributional fluke, and instead a pattern that could carry strategic, or even political, significance. (Levon 2010: 151)

Even though there is a distributional difference in the use of gender-specific morphology between men and women, neither the category 'sex' nor the category 'sexuality' are sufficient for an adequate explanation of the participants' language use. Levon argues that politically active lesbians – "Radical women" in his terms – made use of gender-specific morphology to express their political convictions:

> When referring to women with whom they affiliate or identify, the Radical women use feminine gender morphology to a significantly greater extent than when referring to women they deem to be 'out-group' (in which case they tend to use generic masculine forms). (Levon 2010: 164)

Based on his fieldwork observations and the analysis of several linguistic variables (variation in mean pitch besides variation in the use of gender-specific morphology), he concludes:

> I argue that it is the conjunction of sexual political and, at times, gender identifications that determine how speakers imagine and linguistically constitute their sexualities. The identification that seems to influence speakers' use of language the most is affiliation with a political institution (in a technical, sociological sense)[.] (Levon 2010: 164)

Furthermore, he indirectly criticizes the premises of first and second wave studies by claiming that investigations of "normal categories of sexuality, class, ethnicity, etc." are inadequate to "depict the ways in which 'identity' is experienced by people in their daily lives" (Levon 2010: 166).

2.1.2.5 Summary

We have seen how studies on linguistic variation have been developing over more than five decades. This development is characterized by the narrowing the focus from concepts like 'social class' to individually performed processes of identity construction. Complex concepts such as 'gender' and 'ethnicity,' which are not easily operationalizable, have been found to be relevant for the study of linguistic variation. Consequently, the outlook on the function of linguistic variation has changed from the notion of variants as indexical markers of social categories to expressions of processes of identity construction which can only be deciphered contextually and considering the speakers' agency.

From this theoretically updated perspective, any analytic category which is applied by researchers to the researched population has to be questioned explicitly and needs to be accounted for in the research design. Barron & Schneider propose an altogether different treatment of macro-sociological factors for linguistic variation, based on the equal treatment of 5 – 7 factors, which implies downgrading 'regionality' to a social construct:

> Regarding region, we can, by analogy [to gender and ethnicity], say that we are not interested in geographical facts, but in regional affiliations and identities as they manifest themselves in language use [...] (Barron & Schneider 2009: 427)

In the outline of their framework 'variational pragmatics,' they advocate for a focus on "one macro-social factor at a time," until a better understanding of the interplay

of the factors can be reached (Schneider & Barron 2008: 19). While it is clear that the complexity of the interplaying factors is a methodological challenge, it seems counter-intuitive to recede to an isolated treatment of these factors – especially of 'regionality' – as they suggest. More promising is an approach which focuses on the analysis of the speakers' institutions, like the different activist groups in Levon's (2010) study. As Berger & Luckmann emphasize, any institutional order has to be studied via the knowledge of its members:

> The analysis of roles is of particular importance to the sociology of knowledge because it reveals the mediations between the macroscopic universes of meaning objectivated in a society and the ways by which these universes are subjectively real to individuals. (Berger & Luckmann 1967: 96)

Projected on third wave studies, it is an analysis of 'in-vivo codes' (in GTM terminology), just as "jocks" and "burn-outs:" the categories which are used by Eckert (1989) in her study on the social structure of students in a Michigan high-school. Because of the irreducible complexity of the interdependent concepts, a sound foundation in sociological theory and the explication of theoretic premises is inevitable for any qualitative study of linguistic variation.

2.1.3 *Varietätenlinguistik*

Chambers & Trudgill (1998: 20) argue that dialectology was revitalized in the late twentieth century, due to new technological possibilities for the handling of large amounts of data and due to the theoretical input which was generated by studies on sociolinguistic variation. European dialectologists were looking for ways to include other factors besides 'regionality' in their framework, which led to a shift to research on 'language varieties' – *Varietäten* in German. Still, most approaches within this research paradigm have been foregrounding regional differences because *Varietätenlinguistik* is essentially based on dialectology (Sinner 2014: 24 and Chambers & Trudgill 1998: 21). Just as Labovian studies on sociolinguistic variation, *Varietätenlinguistik* aimed to legitimize the empirical status of the vernacular language and promoted a usage-based study of linguistic variation. For a better distinction between the two research paradigms, the German term *Varietätenlinguistik* will be used here. Studies on sociolinguistic variation in the Labovian tradition – referred to as *Variationslinguistik* by Krefeld (2015b: 22), have been focusing on linguistic variants and the conditions of their usage (see above 2.1.2.1). *Varietätenlinguistik* operates on a more abstract level with the aim to describe clusters of variants as 'varieties' (Krefeld 2015a: 394).

This research paradigm can be traced back to Flydal and Coseriu, who in turn built on Saussure's structuralist theory (Krefeld 2018: 1, Bokelmann 2020: 27 and Sinner 2014: 64). Coseriu (1973: 38–39) established the distinction between 'diatopic' (based on regional factors), 'diastratic' (based on socio-cultural stratification) and 'diaphasic' (based on different styles or registers) differences within a language. This terminology has been influential in philology and especially in Germany based Romanistics (Sinner 2014: 63, 68). Koch (2003: 105) adapted this terminology and added a fourth 'diamesic' dimension which relates to differences between communicative proximity and distance. These differences are understood as the prototypical opposition of spoken and written language. In Koch's model, diastratic is intended to refer both to variation between social layers (vertically) and between social groups (horizontally) (Koch 2003: 103).

This terminology was intended for the explanation of systematic relations between popular notions such as '(standard) language,' 'dialects' and 'sociolects.' Ammon (1987: 317) defines languages as "sets of varieties" and varieties as "elements of languages," whereas "standard varieties and dialects (= dialectal varieties) [...] are various types of such elements (varieties)."

Central to this approach is the equal treatment of language varieties in respect to their functionality: Srhir (2016: 23) argues that all varieties equally serve the communicative and social needs of their speakers. This concept of language varieties is comparable to Luhmann's notion of autopoietic systems: varieties are seen as autonomous systems within the linguistic system. Although Luhmann himself does not treat language as a system on its own, but as "structure or medium" (Bülow 2017: 91–92).[10]

In *Varietätenlinguistik*, top-down classifications have been applied by researchers on the basis of the co-occurring variants, mainly based on phonetic and lexical phenomena (cf. Krefeld 2015b: 2). Recordings of participants' speech along specific variables, in certain regions or within certain social groups, were used to analyze linguistic variants, with the aim to "discover" varieties (cf. Sinner 2014: 69).

Krefeld argues that this approach has two fundamental weaknesses: the first problem is the conceptualization of diatopic, diastratic and diaphasic varieties as equal entities. At least in Europe, diatopic variation is the most evident form of covariation between linguistic and extra-linguistic variables – in this case 'regionality' (Krefeld & Pustka 2010: 18). Therefore, classifications of diatopic varieties – the subject of dialectology – can be seen as unproblematic in a European context (Krefeld 2015a: 396). Whereas regionally defined dialects are understood as prototypical varieties because of their complete functional autonomy, the conceptualization of sociolects

---

10 My translation from the German original.

as complete and functionally autonomous systems is questionable (Krefeld 2015a: 394). It is not surprising that an asymmetric conceptual treatment of the notions 'variety' and 'standard' or 'dialect' is perpetuated by some researchers.

> [Language] is thus schematic for its instances, and 'speaking a language' invariably implies speaking a given variety of that language. However, there is still a widespread tendency to apply a model according to which the standard variety equals the language that it forms part of, and that model is not only at work in folk perception, but also in many branches of linguistics. The impression that we are working at the level of langue [...] when standard varieties form the basis of our analysis, but at the level of parole when the object of study is a nonstandard variety, is obviously misleading, if both of them are just that: varieties of the same language. (Kristiansen 2008: 58)

If, in a European context, dialects are understood as ideal types of varieties, how are ideas of an ideal variety shaped in societies where the factor 'regionality' is not believed to affect language use in the way it does in Europe? If the use of MH in Israel is not conditioned by regional differences, which factors determine the ideal type of a Hebrew variety?

In the previous section (2.1.2.5), I argued that treating factors like 'regionality' as absolute categories cannot adequately explain linguistic variation. Consequently, the conceptualization of diatopic, diastratic and diaphasic variation as neatly distinguishable entities is inadequate as well. Instead, these factors need to be analyzed in conjunction with the speakers' processes of identity construction. This leads us to Krefeld's second point of critique about the traditional focus of *Varietätenlinguistik* on production data (PD). He formulates an updated agenda for the framework, which resonates with third wave variationist theory:

> In linguistics, procedural knowledge is still very much foregrounded because relevant data is mostly and in many cases exclusively obtained from utterances, i.e. from speech production. However, declarative knowledge is extremely important for capturing variation: In a certain sense, *Variations-* and *Varietätenlinguistik* essentially elevate the speakers' procedural and declarative knowledge to the linguist's epistemological level.[11] (Krefeld 2015a: 398, my translation)

As Berruto (1987: 264) noted, any definition of a 'language variety' which is based on structural criteria, like the systematic co-ocurrence of several variants, remains

---

11 German original: In der Sprachwissenschaft steht das prozedurale Wissen immer noch stark im Vordergrund, denn die relevanten Daten werden meistens und in vielen Fällen ausschließlich aus Äußerungen, d.h. aus der Sprachproduktion gewonnen. Das deklarative Wissen ist jedoch äußerst wichtig für die Erfassung der Variation: In gewisser Hinsicht sind Variations- und Varietätenlinguistik im Kern nichts anderes als die Hebung des prozeduralen und deklarativen Sprecherwissens auf die epistemologische Ebene des Sprachwissenschaftlers.

unsatisfactory. Therefore it has been argued, that a definition of 'variety' taking into account the speaker's knowledge is more sensible (cf. Sinner 2014: 20). Krefeld (2018: 2) specifies that it is outright wrong if linguists classify utterances as being marked in a certain way and thus belonging to a certain variety because this type of information cannot be extracted from PD alone. To derive meaningful classifications, the speakers' 'declarative knowledge' (DK) about linguistic variation has to be taken into account (Krefeld 2019: 28). Furthermore, PD and DK have to be analyzed in conjunction with *Perzeptionsdaten* 'perceptual data' which leads to a methodological realignment of the research paradigm under the new label *Perzeptive Varietätenlinguistik* (Krefeld & Pustka 2010).

2.1.4 *Perzeptive Varietätenlinguistik* (PVL)

Perzeptive Varietätenlinguistik (PVL) brings the speakers back into the picture, by relating their representations of variation with their perception of variants.[12] The speaker himself is regarded as the primary agent who causes variation:

> The locus of variation is not an abstract point in a matrix of formal parameters, but a speaker in his historical concretion: with his repertoire of varieties (of more or less languages), with the communicative routines that regulate his use of available varieties in the networks of his communication partners, and with the repercussions of these routines on his own repertoire.[13] (Krefeld 2015a: 401, my translation)

While *Varietätenlinguistik* builds on structuralism, just as the first and second wave studies on sociolinguistic variation, PVL is conceptually close to the third wave of variationist studies, with its post-structuralist perspective on the individual speakers and their construction of meaning.

In their model, which is the basis for PVL, Krefeld & Pustka (Forthcoming) elaborate on the relation between the speakers' knowledge and behavior (see Fig. 2.2). They argue for a clear conceptual distinction between representation and perception of linguistic phenomena: While perception is confined to a real usage event, representation are part of the speakers' knowledge about language. Representations can

---

12 An English translation of PVL is not yet established. I suggest to translate the original German term as 'Perceptual Variationist Linguistics.'

13 German original: Der Ort der Variation ist kein abstrakter Punkt in einer Matrix formaler Parameter, sondern ein Sprecher in seiner historischen Konkretion: mit seinem Repertoire an Varietäten (von mehr oder weniger Sprachen), mit den kommunikativen Routinen, die den Gebrauch der ihm verfügbaren Varietäten in den Netzwerken seiner Kommunikationspartner regeln, und mit den Rückwirkungen dieser Routinen auf sein eigenes Repertoire.

be elicited independently, without the context of actual language usage and without any perceptual basis. According to Krefeld & Pustka (2010: 14), PVL is a sub-discipline of perceptual dialectology or folk linguistics as popularized by Niedzielski & Preston (2003). Like third wave studies (see 2.1.2.2), perceptual dialectology is situated within the theoretic context of the sociology of knowledge.

> The sociology of knowledge [doesn't seek] new apriorisms or hypothetical-deductive systems for deriving and testing all knowledge, but metatheories in a quite different sense: namely downstream theories that explain how knowledge that discriminates with respect to objects (and in this sense claims to be true) arises and is maintained in social contexts.[14] (Luhmann 1993: 61, my translation)

In analogy to Luhmann's above definition, perceptual dialectology and PVL are linguistic meta-theories to study knowledge about language and linguistic variation, its emergence and its fostering through conversation. The discriminating aspect about this knowledge (in Luhmann's words) is related to the cultivation of linguistic norms which define 'correct language' (see also 2.1.2.3). Luhmann (1993: 61) claims that these "metatheories can be empirically operationalized," which is the aim of PVL.[15]

In the context of PVL, this means that perception and representation have to be treated separately and that representations alone cannot be understood as a proof for the existence of varieties or their meaning. In the next section, theoretic premises from PVL for the organization of linguistic knowledge will be revised and terminologically defined for further use.

2.1.4.1 The Organization of knowledge about language and linguistic variation

From our understanding of linguistic variation (see 2.1.2.2) follows that a purely structuralist outlook on linguistic knowledge, as a symbolic pairing of form and function, is too simplistic. Schmid provides a more encompassing perspective on the linguistic sign by defining 'utterance types' as follows:

> [U]tterance types do not just consist of pairings of form and meanings and vice versa, but include information about communicative goals (What are the participants trying to achieve?), the diverse usage activities (What are they actually doing?), and in which contexts the usage

---

14 German original: Die Wissenssoziologie [sucht] nicht neue Apriorismen oder hypothetisch-deduktive Systeme zu Ableitung und Überprüfung allen Wissens, sondern Metatheorien in einem ganz anderen Sinne; nämlich nachgeschaltete Theorien, die erklären, wie Wissen, das in bezug auf Gegenstände diskriminiert (und in diesem Sinne wahr zu sein beansprucht) in sozialen Kontexten entsteht und gepflegt wird[.]
15 My translation from the German original.

event is taking place (Who is speaking to whom in what physical and social situation?). [...] [C]onventionalized utterance types can hence be regarded as condensed records of their own usage history. [...] Without this information, linguistic conventions and linguistic knowledge would remain crude and lack the subtlety they have with native speakers, as opposed to speech typical of someone who begins to learn a foreign language. (Schmid 2020: 16)

The type of information beyond form and meaning that speakers associate with utterance types is the missing piece of the puzzle which I was looking for as a Hebrew learner (cf. 1). This study is an investigation about this type of knowledge by the means of the analysis of HSs' representations of linguistic variation.

According to Krefeld (2015a: 397-398), the speakers' knowledge about language is of procedural and of declarative nature: procedural knowledge is a premise for the ability to speak and to use certain variants, while declarative knowledge (DK) contains representations which are associated with the variants. In Schmid's (2020: 27) cognitively oriented terminology, the equivalent of procedural knowledge, are "conventionalized utterance types" and he terms "their mental representations as entrenched patterns of associations." In the following, Krefeld's terminology will be used because this study is situated within a variationist framework.

The speakers' DK is understood to contain information about the markedness of linguistic forms (Krefeld & Pustka 2010: 12). In general, 'markedness' is a concept which is used to describe "certain types of asymmetries within categories" (Lakoff 1987: 59). For example, a linguistic form can be marked as a regional variant and, depending on the context, it can be judged as 'incorrect.' The so-profiled representation of a linguistic form is a premise for its perception as being "marked in a certain way," when the linguistic form is encountered in an actual usage event. Without the representation of the form as marked, it will be perceived as inconspicuous (Krefeld 2015a: 398). As has been stressed in 2.1.2.2, markedness is not an inherent quality of a linguistic form, but dependent on the context of its usage and the agency of the speaker.

This general distinction of marked and unmarked linguistic forms is the basis for the distinction of 'standard' and 'non-standard variants' which lies at the heart of variationist linguistics. It is the premise for linguistic indexicality (Silverstein 2003): the symbolic link between a variant and a concept – such as the indexical relation between *sibolet* and the Ephraimites (see 2.1 above). PVL defines 'standard,' in contrast to the 'normative correct language,' as the unmarked (for its speakers) variety of a language

It should be noted that for the vast majority of those who use it as a matter of course, so to speak by default, the standard actually provides the neutral background against which other

variants and varieties stand out as salient and in this respect should be considered as marked.[16]
(Krefeld 2011: 104, my translation)

This understanding is based on the psychological model of "figure/ground segrega-
tion" (Ungerer & Schmid 2006: 163; see 2.1.4.3): All variation can only be perceived in
relation to a neutral background or 'linguistic standard.' In contrast, 'non-standard
varieties' can be modeled as being perceived by the speakers as a complex pattern
– a *Gestalt*, of co-occurring marked variants (Krefeld 2015b: 23). Typically, marked
varieties are associated with marked identities (Buchholtz & Hall 2004: 372–373).
Based on this definition, our notion of linguistic norms differs from explicitly fixed
grammatical rules and extends to the domain of pragmatics i.e. "how language is used
to communicate" (Levinson 2024: 1). Discourses about flexible conventions which
affect language use in circumstantially changing usage events are the research object
of metapragmatics (Silverstein 2003: 196 and Spitzmüller 2019).

From the perspective of the language learner, it is easy to see that procedural
knowledge of how to use a certain variant is ultimately dependent on DK. Language
learners who have mastered grammatical rules and textbooks, but have missed
out on the actual usage of the language, are essentially lacking DK. Their lack of
participation in the speakers' stock of knowledge can cause difficulties in social
interaction.

Participation in the social stock of knowledge thus permits the 'location' of individuals in society
and the 'handling' of them in the appropriate manner. This is not possible for who does not
participate in this knowledge, such as a foreigner [...] (Berger & Luckmann 1967: 56)

In 2.1.4.4 the "location" and the "handling" of speakers during a usage event will be
reassessed theoretically. The next section is a short excursion about the necessity to
integrate different perspectives on the research subject for a sound PVL methodology
and metapragmatic research. Then, the scope of the term 'representation' will be
reassessed in 2.1.4.3 for a theoretic perspective on the organization of knowledge
about language and linguistic variation.

2.1.4.2 A question of perspective

A general distinction of possible perspectives on the research subject is helpful for
its adequate treatment. Therefore, Krefeld & Pustka (2010: 23) reassess the relations

---

16 German original: Festzuhalten ist [...], dass der Standard für die große Masse derjenigen, die sich
seiner mit schlichter Selbstverständlichkeit, so zusagen *by default*, ständig bedienen tatsächlich den
neutralen Hintergrund liefert, vor dem sich andere Varianten und Varietäten als salient hervorheben
und insofern als markiert zu betrachten sind.

of four epistemic dimensions which are defined by a twofold distinction between an 'emic' and an 'etic'perspective and between an expert (e.g. a linguist) and a lay (e.g. a speaker) perspective. The distinction of 'emic' and 'etic' was introduced by Pike (2015 [1967]) and expresses the opposition of categories which can either be structurally defined from within the system (of a specific language), just as phon*emic* categories, or independently from the system, just as phon*etic* categories. PVL upholds that valid insights into the nature and the dynamics of linguistic variation can only be reached in the light of the speakers' representations and perceptions thereof. Therefore, any of the four perspectives, on its own, is insufficient for the aim of PVL.

An expert-etic perspective can lead to the postulation of varieties while ignoring the speakers' perceptions. A fully speaker-emic perspective leads to an unscientific treatment of representations as facts. In this context, it is easily ignored that linguists are speakers of languages and therefore, subject to the emic perspective, as well. Their expert status does not change the nature of the representations about their spoken language – they are representations all the same. Often the collection of DK is not well explicated because the researcher is part of or at least very close to the researched population (and subject to a lay-emic perspective): it is supposed, not only, that her or his representations of social and linguistic behavior is shared by the researched population – but, that her or his interpretation of its meaning is shared by the readers as well. In research on MH varieties, this shortcoming of explication and contextualization of the researchers' background is widespread. Most researchers in this field are L1 Hebrew speakers and tend to present their individual representations as common sense – without empirical basis from other Hebrew speakers. This lack of explication of the researchers' hypothesis can render research incommensurable for cross-cultural comparisons and typological aims.

From this respect, it can be beneficial for the study to occupy an outsider position as researcher which forces oneself to a more conscious approach to the researched population, with less implicit knowledge at hand. In consequence, readers who are unfamiliar to the context of the research can be involved more easily. However, there is no ideal solution to this theoretic problem of perspectives. As will be argued in 2.2.1, a conscious approach to the field, openness and a detailed explication of the research methodology can improve the overall quality of research.

2.1.4.3 Perception, representation and concepts

The term 'representation' has a long history in philosophy and has become a central concept in cognitive science, although its usage remains somewhat ambiguous (Sinha 2007: 1280). Berger & Luckmann (1967: 223) use representation in "the Durkheimian usage, but in broader scope", primarily in describing the "representation of an institution in and by roles" as "the representation *par excellence*" (Berger & Luckmann

1967: 93). Krefeld & Pustka (2010: 11) base their notion of representation on its usage for the organization of knowledge in psychology, in social science (in reference to Durkheim and Bourdieu) and to structuralist linguistics with its idealized notions of phonological, syntactic and semantic representations – what they term linguistic representations "in the narrow sense" (Krefeld & Pustka Forthcoming).

In cognitive linguistics, linguistic structure is studied to gain insights into cognitive processes. Linguistic structure is understood as being "motivated by conceptual representation and communicative function" (Sinha 2007: 1280). Because humans can only perceive the world through bodily structures, 'embodiment' is a key principle affecting the processes of human categorization. For example, events are typically perceived as made up of structured and wholesome entities (*gestalt*) with typified relations. Figure 2.1 is a simple schema of the perceptual process of a spatial relation.
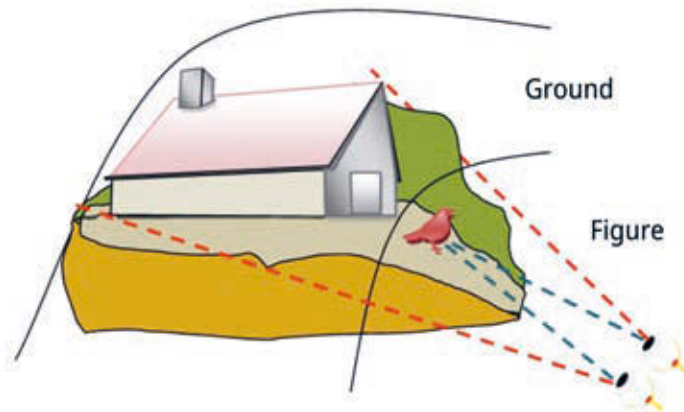

Fig. 2.1: Perception of a spatial relation (reproduced from Schulze 2015)

Evans (2010: 21–22) describes the process of perception as the transformation of external sensory information, such as light, into a "perceptual object." Thus, the situation in Figure 2.1 is at first organized into a smaller and a bigger perceptual object which are indicated by the blue and the red dashed lines. The smaller and potentially moveable perceptual object is usually referred to as 'figure,' whereas the bigger and typically static perceptual object is referred to as 'ground' (Talmy 2000: 184, Ungerer & Schmid 2006: 164). These perceptual objects are then bestowed with meaning by the application of the concepts 'bird' and 'house.' Evans distinguishes 'perceptual objects,' which she terms "percepts," and 'concepts' as follows:

> Percepts are typically available to conscious experience. That is, they are the product of on-line processing, resulting from a stimulus array perceived in the 'here-and-now'. A consequence of this is that they consist of specific information relating to the specific stimulus array that they are derived from. Thus, they are episodic in nature. Concepts, on the other hand, represent schematisations, formed by abstracting away points of differences in order to produce representations which generalise over points of similarity. [...] concepts are representations in the sense of re-presentations. That is, they are stored in memory and can be activated during off-line processing. That is, they can be recalled in the absence of the percept(s) which may have given rise to them. (Evans 2010: 21–22)

The relation between the processes of representation and perception is again described in a more general definition:

> [R]epresentations represent something. We ordinarily think of both our representational mental states, such as beliefs and desires, and external representational artifacts, such as maps and words, as being about some object, property, or state of affairs. [...] Second, while this phenomenon of aboutness or intentional directedness seems to involve a kind of relation, if so, it must be a very special kind of relation, in which the distal relatum – the intentional object of the representation – needn't actually exist. To put it less tendentiously, it seems central to our conception of representations that they can misrepresent. Finally, we ordinarily think of the content of a representation as being somehow relevant to the causal role that the representation plays. (Morgan 2014: 218)

It is important to note that perceptual objects cannot be identical with any form of representation – even though they may function as the experiential basis for the representation. The relation between the processes of representation and perception is intricate because they are interdependent only to some degree – their relation is not necessarily causal (Matsumoto-Gray 2009: 114). Having just touched upon this complex philosophical topic which is subsumed under the label 'relativism,' it should have become clear that the conflation of perception and representation is problematic.

While Kehrein et al. (2010: 380–381) distinguish between "Perzeption" and "Projektion" in their model of conceptualization, they also use "Representation" for both processes interchangeably. What they term as "Perzeption" is the formation of a representation based on an actual experience and "Projektion" is understood as the complementary process of applying a representation onto an experience. This distinction is important because it highlights the possibility of misrepresentation by projecting an inadequate concept onto a perceptual object. Misrepresentations of linguistic variation can lead speakers to the wrong classification of production data which can be witnessed during perception experiments (Kehrein et al. 2010: 378–379). For the purpose of PVL, Krefeld & Pustka (Forthcoming) stress that "we have to distinguish between representations which are based on perception and

those which are based on other experiences in the non-linguistic world, so-called 'pseudo-linguistic knowledge' or 'linguistic myths'."
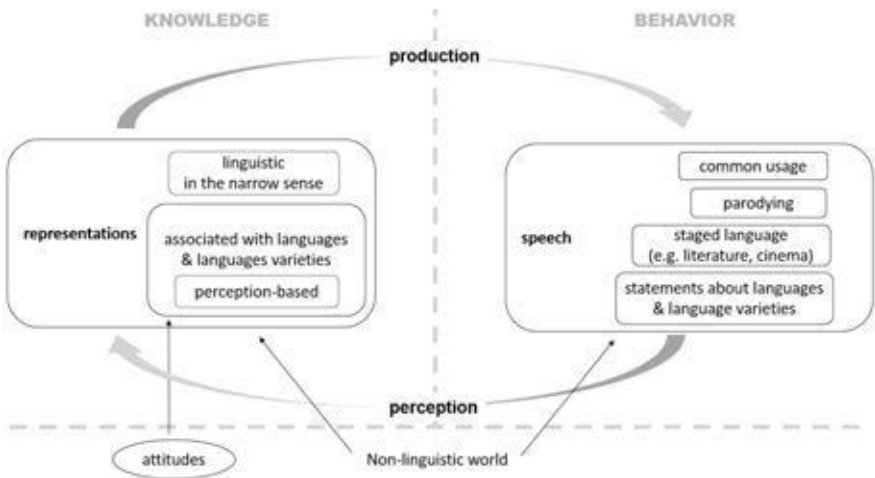

Fig. 2.2: Speaker's knowledge and behavior (reproduced from Krefeld & Pustka Forthcoming)

Krefeld & Pustka's (Forthcoming) model which is reproduced in Fig. 2.2 is an updated version of their earlier model (see Krefeld & Pustka 2010: 12). It displays the intricate relations between speech, its representations and the processes of perception and production, along the dimensions of behavior and knowledge. Basically, Krefeld & Pustka (Forthcoming) posit with the model that not only linguistic phenomena are represented as linguistic categories like 'phonemes' – which is referred to by "linguistic in the narrow sense" – but that languages and language varieties are equally represented in the speakers' DK. The model illustrates that the formation of representations about linguistic variation is potentially influenced by the perception of a usage event. Additionally, extra-linguistic factors influence the process of representation and therefore, the speakers' statements about language and linguistic variation. Kehrein et al. render a vivid description of the speakers' DK and its dependence on extra-linguistic factors:

> This knowledge is to be understood as part of an individual's linguistic and world knowledge and thus as a complex mental structure that is influenced by all aspects of everyday life experiences (travel, communication, media, prejudices , etc.), both linguistic and non-linguistic. Consequently, this mental structure cannot be derived from linguistic phenomena alone, but is constituted individually from the totality of possible experiential content – it defines the frame-

work for (linguistic) interaction in the form of individually constructed everyday categories (=concepts).[17] (Kehrein et al. 2010: 352, my translation)

In fact, representations of linguistic variation can be formed and applied independently of any perceptual basis (Krefeld & Pustka 2010: 12 and Kehrein et al. 2010: 378–379). In this context, it may be helpful to reassess Berger & Luckmann's definition of reification for a better understanding of the processes causing "pseudo-linguistic knowledge" or the misrepresentation of linguistic phenomena:

> [R]eification is the apprehension of the products of human activity as if they were something other than human products – such as facts of nature, results of cosmic laws, or manifestations of divine will. Reification implies that man is capable of forgetting his own authorship of the human world, and, further, that the dialectic between man, the producer, and his products is lost to consciousness. [...] reification can be described as an extreme step in the process of objectivation, whereby the objectivated world loses its comprehensibility as a human enterprise and becomes fixated as a non-human, non-humanizable, inert facticity. (Berger & Luckmann 1967: 106)

I argued that concepts such as 'correct Hebrew' can be understood as institutions in Berger & Luckmann's (1967) terms (see 2.1.2.3). Because institutions and roles can be reified, concepts such as 'Oriental dialect of Hebrew' can also be apprehended as facticity – regardless of any perceptual basis. As for any normative standard of a language, it is probably impossible to find a sound perceptual basis for the concept 'correct Hebrew' because there is hardly any HS who consequently adheres to all normative rules which define 'correct Hebrew' – therefore it is a reified concept. Interestingly, Berger & Luckmann (1967: 108–109) use 'Jewishness' as an example for reification. In our context, the conceptualization of Hebrew as a 'Jewish language' takes this reification one step further.

Again, these aspects are summarized under the notion 'linguistic stereotype'

> Humans have receptive competence of lectal varieties, but the images formed are not necessarily accurate, at least from the perspective of experts such as linguists. It is in this sense that I use the term "linguistic stereotype"; an instance of folk perception with respect to the distinctive features of a speech variety. Linguistic stereotypes, rather than representing exaggerated and distorted images, constitute useful cognitive reference point constructions which allow

---

17 German original: Dieses Wissen ist als Teil des Sprach- und Weltwissens eines Individuums und demnach als komplexe mentale Struktur zu verstehen, die von allen – sprachlichen wie nichtsprachlichen – Aspekten des alltäglichen (Er-)Lebens beeinflusst wird (Reisen, Kommunikation, Medien, Vorurteile etc.) Insofern ist diese mentale Struktur nicht allein aus sprachlichen Phänomenen ableitbar, sondern konstituiert sich aus der Gesamtheit der möglichen Erfahrungsinhalte individuell und definiert in Form von individuell konstruierten Alltagskategorien (=Konzepten) den Rahmen für (sprachliche) Interaktion.

> us not only to categorize our linguistic environment but also to categorize and characterize our social environment. As I have previously argued (Kristiansen 2003), linguistic stereotypes metonymically evoke the social stereotypes associated with the group in question. Accents are thus not only regionally and socially diagnostic; they also serve to characterize speakers in very significant ways. (Kristiansen 2008: 61)

Rather than just highlighting the artificial nature of linguistic stereotypes, the importance of these representations for the speakers' processes of categorization is expressed in the second part of the citation. Agha (2006: 148) similarly defines "metapragmatic stereotypes" as "culture-internal models of utterance indexicality associated with speech variants." The study of these types of knowledge as the representations of speakers who act as research participants is central to PVL (Krefeld 2019: 28).

As Kristiansen (2008: 53) points out, "from a Cognitive Linguistics perspective, the link between lectal and social categories is a complex and intriguing affair." This topic is addressed in this study by RQ8 and RQ9. I will try to approach these questions by identifying and analyzing the most commonly used categories of this type from the participants' statements. Based on the theoretic discussion so far, I argue that these categories are shaped according to general principles of natural categorization, as they are studied in cognitive linguistics.

To determine the content of these categories, it is necessary to study what is termed "attitudes" in Krefeld & Pustka's (Forthcoming) model – this research goal is addressed by RQ6. I will assume that frequently expressed LAs hint at common representations. LAs can be grasped as integral components in the construction of meaning and the social function of language. For example, the metaphorical concept LANGUAGE AS A BOND can be used to serve these processes when defining Hebrew as a 'Jewish language' (cf. Berthele 2008: 309). In the following, LAs are understood as constructed in relation to concepts which have been reviewed, so far: institutions, roles and representations of linguistic variation. Just as the markedness of a linguistic variant is sensitive to the context and subject to diachronic change, LAs are not fixed categories which can be linked inseparably to a participant, once they have been expressed. For instance, participants expressed different and at times contradicting attitudes towards 'slang' or other linguistic concepts during interviews for this study, in the course of less than one hour.

2.1.4.4 Categorizing usage events

From a structuralist perspective, language varieties are understood as equally functional for the exchange of information. However, linguistic variation is far more than redundant structure: each variant is associated with extra-linguistic parameters that fulfill different pragmatic functions (Bokelmann 2020: 38, Krefeld 2019: 27). Language

is used to order the environment and to make sense of it through categorization. Thus, speakers use language to convey meaning to their own actions by performing their identity and demonstrating their belonging to groups (see 2.1.2.2). Because acts of identity construction are essentially related to in-group and out-group differences and performed through accommodation during usage events, the classification of other speakers is an important part of these processes (Kristiansen 2008: 73–74). There are not just categories that are inherently used in each language to bestow order and meaning. At the same time, linguistic phenomena and their speakers are subject to further categorization.

> When we meet people for the first time, engage in the maintenance of social relationships or the negotiation of social roles and positions, the so-called ideational, message-bearing or referential function of language is secondary with respect to the so-called social, phatic or interpersonal function. Language is not just a tool for ideational communication. It is also an important vehicle for social communication. It enables us to identify and characterize unknown individuals, convey and preserve our own relative position on a given hierarchical scale and define ourselves as ingroup or outgroup members of relevant social categorizations. (Kristiansen 2008: 70)

This process of classification of other speakers during a usage event which was reassessed above with Berger & Luckmann's (1967: 42) terms as the "location" of speakers "in society" is also described by r36f3l1's statement:

(2) r36f3l1 (15:30)

There is this automatic thought to divide people somehow. Let's say I'm in Jerusalem – so I speak with someone and I think instantly whether he is Jew or Arab. Then, I hear his accent – I understand that he is Arab.

יש כזאת מחשבה אוטומטית לחלק אנשים איכשהו. נגיד אני בירושלים אז אני מדברת עם מישהו ואני ישר חושבת אם הוא יהודי או ערבי. ואני שומעת המבטא שלו, אני מבינה שהוא ערבי.

R36f3l1 describes her thought pattern as "automatic." It can be inferred that it would be hard for her to stop these thoughts of how to "divide people." This involuntary and at the same time necessary aspect about the categorization of people along stereotypes is also described by Nassehi:

> We can hardly move around in our everyday lives without stereotyping ideas of order. To infer role expectations from partial aspects of people's behaviour, we must bear in mind typologies and taxonomies that make us capable of acting at all. We can assess what the other will do and what we might expect from that person only if we are in a position to draw on social structures in the form of types. (Nassehi 2024: 33–34)

The type of "lectal categorization," in Kristiansen's (2008: 73) terminology, that r36f3l1 explains is the classification of someone in Jerusalem as "Arab" on the basis of his "accent" – it can be inferred that she operates with a category 'Arabs.' Kristiansen renders a general theoretic account of this type of categorization processes, using cognitive linguistics terminology:

> We assume that Hearer categorizes the speech pattern of Speaker and evokes an entrenched, metonymic schema concerning the speech style in question and the social categorization that effected it. The central images of both lectal and social categorizations operate as cognitive reference point-constructions. Such categories are presumably naturally organised around clusters of elements in the visual or auditory modalities which are distinctive enough to establish perceptual contrast with neighbouring categories. Lectal categorization thus seems to involve a conceptualizer who correlates a token (stretch of unidentified speech) with a number of idealized speech models (linguistic stereotypes). (Kristiansen 2008: 73)

This type of social and linguistic categorization processes can also be grasped in terms of Harvey Sacks' "membership categorization device" (MCD):

> A Membership Categorization Device is composed of two parts – first, one or more collection(s) of categories, and, second, some rules of application. [...] The categories of person (or member [of the society] in Sacks' parlance) which figure in interaction and in social life more generally are not a simple, single aggregate of categories, but are organized into collections of categories. A collection is a set of categories that 'go together' – for example, [male/female]; [Buddhist/Catholic/Jew/Muslim/Protestant . . .] (Schegloff 2007: 467)

R36f3l1's statement does not include any information about other categories, besides 'Arabs,' which she might use. Therefore, we can only infer that she disposes of a collection of categories which contains at least [in-group/Arab].

Schegloff characterizes the categories of a MCD as

> the store house and the filing system for the common-sense knowledge that ordinary people – that means ALL people in their capacity as ordinary people – have about what people are like, how they behave, etc. (Schegloff 2007: 469)

He goes on to define some common properties of the MCD categories, such as "category-bound activities" that "are kinds of activities or actions or forms of conduct taken by the common-sense or vernacular culture to be specially characteristic of a category's members" (Schegloff 2007: 470). In r36f3l1's example (2), having an "accent" can be interpreted as a category-bound activity belonging to her category 'Arabs.' In Sacks's (1972: 35) original paper, the example which he uses as a category-bound activity of 'babies' is 'crying.'

From a cognitive scientific perspective, it may well be that MCD categories come close to what is discussed in the literature as 'basic level categories' (see

Murphy 2002 for an overview). Basic level cagetories (BLC) are defined by their association with prototypical activities, among other properties (Schmid 1996: 292). The MCD categories are abstract concepts and cannot be measured with scientific means – unlike colors which can be defined by the wave length of light. Even though most research revolving around the concept of BLCs has been carried out with measurable objects, there is no reason why a general definition should not apply to the categorization of linguistic variation, as well:

> Basicness in categorization has to do with matters of human psychology: ease of perception, memory, learning, naming and use. Basicness of level has no objective status external to human beings. (Lakoff 1987: 38)

During a usage event which takes place as face-to-face encounter, an individual's visual appearance and his way of speaking are probably the most easily perceivable characteristics. Accents certainly qualify as actions which are typically learned together with category names for their speakers. Stereotypical accents are used, for example, in movies to portray characters as belonging to certain categories. This aspect is taken up in Krefeld & Pustka's (Forthcoming) model as "parodying" (see Fig. 2.2). When perceiving linguistic variation, for example, in the form of someone who speaks Hebrew with an 'accent,' it is very likely that the perceiver adopts his behavior accordingly (Kristiansen 2008: 73–74): Many people switch to a simplified language variety or another language, when they perceive a 'foreign accent' (see also the account of my own experience in 1). These acts of accommodation can be understood as prototypical (re)actions which are associated with categories for linguistic variation. "[B]asicness of naming and use," in Lakoff's (1987: 38) terms, is expressed in Sack's "economy rule" with which he defines MCD categories:

> It holds that a single category term from any MCD can in principle do adequate reference. More can be used; as for example in a reference to a '45-year-old Russian ballerina'; but, in principle, one term can do adequate reference. (Schegloff 2007: 471)

In confirmation of Kristiansen's (2008: 73) assertion (see above) further characteristics of natural categories, as they are described by Rosch (1978) seem to apply. I will assume that representations of linguistic variation and the speakers are structured according to principles which have been established in cognitive linguistics. This argument is elaborated with empirical data in 5 and especially in 5.3.

Another aspect about natural categories that can be illustrated with r36f3l1's statement is their dependence on cultural context:

> It should be noted that the issues in categorization with which we are primarily concerned have to do with explaining the categories found in a culture and coded by the language of that

> culture at a particular point in time. When we speak of the formation of categories, we mean their formation in the culture. (Rosch 1978: 28)

By referring to the the circumstances where she might use the described process of categorization, "Let's say I'm in Jerusalem," r36f3l1 hints to the fact that she uses other collections of categories in other circumstances – for example, in Tel Aviv. Moreover, it seems safe to say that r36f3l1's category 'Arabs' is different from the category 'Arabs' which is used by a Berber in Morocco.

> To categorize a stimulus means to consider it for purposes [...] It is to the organism's advantage not to differentiate one stimulus from others when that differentiation is irrelevant to the purposes at hand. (Rosch 1978: 28–29)

Based on Rosch's (1978: 28–29) above cited principle of "cognitive economy," it is sensible to expect that the number of categories which are conventionally applied by a speech community to categorize linguistic variation is finite and that there has to be some congruity of these categories due to speakers' similar communicative goals. Schegloff (2007: 475) points out that it is intriguing why certain of the many possible categories are used and ultimately, stresses the importance of investigating the categories, as they are used by the speakers themselves:

> The issue is, after all, not whether we can or should make a category out of it, but whether they – the parties to the conversation – do so, and, if they do, what that sounds or looks like. And this, of course, should be sought in data, in an effort to get at what the parties to the talk are doing by talking the way they are. (Schegloff 2007: 477)

This study takes up this research objective by investigating HSs' representations of linguistic variation and social groups.

2.2 The benefits of Grounded Theory Methodology for this study

The empirical investigation of linguistic variation is methodologically challenging because of the theoretical implications that have been discussed so far. How can interpersonal processes for the construction of meaning that are crucial for the analysis of linguistic variation be studied?

Berger & Luckmann (1967: 173) describe the fleeting nature of human experience by referring to Heraclitus' aphorism *panta rhei* 'everything flows' (see 2.1.2.3). In Israel, the metaphor of *lizrom* 'to flow' is used to refer to a flexible and relaxed mindset: because little else can be done in unknown or fast-changing circumstances, the most practicable solution is to go with the flow. The expression *litsnoaḥ ve-lizrom*

'parachute and flow' refers somewhat jokingly to a flexible strategy in response to unknown conditions where one finds oneself after parachuting into enemy territory. This metaphor expresses the conviction that flexibility is a key element for reaching one's goals and for surviving.

In the context of research methodology, a flexible and open mindset of the researcher is viewed as a prerequisite for successful fieldwork and – in general, for any kind of qualitative research. Although a flexible research design is certainly advantageous in an under researched area, some kind of systematization is needed to reach valuable results. Grounded Theory Methodology (GTM) is the systematization of a principally flexible research paradigm. GTM originated in the context of American qualitative sociology, in a climate of academic protest which is also reflected in Labovian variationist studies. The principles of the research paradigm were first published by Glaser & Strauss (1967) who each subsequently developed and propagated their own accentuations of the methodology. GTM is a hermeneutic approach for the sociological analysis of artifacts and processes of symbolization:

> Theoretical concepts and models are developed on the basis of empirical data from everyday contexts, starting from a preliminary problematization perspective, and are constantly and recursively linked back to the experiential level. The corresponding theory of a social section of the world or of a problem is 'grounded.'[18] (Breuer 2010: 39, my translation)

According to Charmaz who worked together with Glaser and Strauss and ultimately propagated her version of 'constructivist grounded theory,'

> Grounded theory demystifies the conduct of qualitative inquiry. Rather than applying a preconceived theoretical framework, your ideas about the data guide how you construct the theoretical analysis. (Charmaz 2004: 54)

GTM is a heterogeneous framework – there are various interpretations by different researchers with a differing degree of detail about its methods. Before discussing the principles of GTM and its benefits for this study, the theoretic foundation of the framework within qualitative sociology is reviewed.

---

18 German original: Auf der Basis von Erfahrungsdaten aus alltagsweltlichen Kontexten werden – von einer vorläufigen Problematisierungsperspektive ausgehend – theoretische Konzepte und Modellierungen entwickelt und dabei fortwährend rekursiv an die Erfahrungsebene zurückgebunden. Die entsprechende Theorie eines sozialen Weltausschnitts bzw. eines Problemthemas wird 'gegenstandsgegründet' herausgearbeitet ('grounded').

2.2.1 Principles of qualitative sociology

> A fundamental principle in qualitative research is appreciating subjective realities and, conversely, rejecting the belief in an objective truth. Every individual perceives the world differently and good research constitutes a comprehensive and complex picture stemming from different perspectives , which will never be free of contradictions.[19] (Dunkelberg 2005: 249, my translation)

The above citation expresses the aim of qualitative research to explore the research area from multiple perspectives. For the qualitatively oriented framework of PVL, the importance of the methodological integration of different perspectives was illustrated in 2.1.4.2. Therefore, a multi-faceted analysis which is potentially contradictory can increase the overall quality of the research by enabling a more enhanced understanding.

Because the concept of scientific objectivity is principally rejected, the researcher's independent analytic position is equally called into question. The role of the researcher is seen as influenced by the constant and paradoxical changing of perspectives between participation in and distancing oneself from the area of research (Bohnsack & Nohl 2001: 32). Thus, the research process is understood ideally as a fruitful interaction between the researcher and the researched population. Instead of testing a preconceived hypothesis on the population within a standardized and inflexible framework, qualitative research promotes co-operation with the participants who are termed *Forschungspartner* 'partners in research' by Breuer (2010: 40) and with other researchers. Subjectivity is no longer perceived as a threat to scientific integrity – as long as it is handled consciously and openly – because it simply cannot be ruled out by any means (Dunkelberg 2005: 250). There is no way to reach universally and objectively valid conclusions by the means of qualitative research:

> Knowledge in Social Science is understood as fundamentally partial and location-bound – it cannot claim general validity, but is limited in its informative value and inevitably linked to the perspective of the researchers.[20] (von Unger 2014b: 22, my translation)

---

19 German original: Ein grundlegendes Prinzip in der qualitativen Forschung ist die Wertschätzung von subjektiven Wirklichkeiten bzw. umgekehrt die Ablehnung des Glaubens an eine objektive Wahrheit. Jedes Individuum nimmt die Welt anders wahr und gute Forschung konstituiert ein umfassendes und komplexes Bild aus verschiedenen Perspektiven, das aber nie widerspruchsfrei sein wird.

20 German original: Sozialwissenschaftliche Erkenntnis wird als grundsätzlich partial und standortgebunden verstanden –- sie kann keine allgemeine Gültigkeit beanspruchen, sondern ist in ihrer Aussagekraft begrenzt und unausweichlich mit der Perspektive der Forschenden verknüpft.

Every qualitative analysis is situated within the specific contexts of its time, its place, researcher(s) and participants. The thorough explication of these contexts and the different perspectives which are inherent in the particular study is regarded as a central requirement for valuable qualitative research:

> Reflexivity regarding the researchers' subjectivity, their positioning in the research field and their influence on the research process is considered a quality feature in qualitative Social Science.[21] (von Unger 2014b: 23, my translation)

Based on these premises, the scientific output of GTM-based research is provisional:

> [T]he published word is not the final one, but only a pause in the never-ending process of generating theory. (Glaser & Strauss 1967: 40)

It is easy for anyone who does not share these theoretic premises to dismiss qualitative social science as merely anecdotal. Nassehi argues against this undue criticism by stressing the core aim of the research paradigm, which is the systematic analysis of supra-individual patterns:

> Strictly speaking, qualitative social research is also a method of pattern recognition that concerns itself with the development of order, that is, with a reconstructible process of ruling out other possibilities. Those who believe that qualitative social research is research that makes the subject speak or that gets closer to reality because the data for research are 'natural' and close to everyday life are simply doing bad sociology, as this kind of research is also about supraindividual patterns and about the methodically controllable recombination of meaning. (Nassehi 2024: 36)

This study adheres to a qualitative research paradigm to recognize supra-individual patterns in participants' statements about linguistic variation in MH, with the aim of analyzing typical representations.

2.2.2 Principles of GTM

Three principles are key to GTM: data-based generation of the hypothesis, a post-structuralist outlook on meaning which leads to a methodological integration of different perspectives on the research area and systematic methodological recursivity. Although the first (usage-basedness) and the second principles are also central to third wave variationist studies and to PVL, GTM has rarely been used as a method-

---

21 German original: (Selbst-) Reflexivität im Hinblick auf die Subjektivität der Forschenden, ihre Positionierung im Forschungsfeld und ihren Einfluss auf den Forschungsprozess gilt in der qualitativen Sozialforschung als Qualitätsmerkmal.

ological basis in linguistic research. Hadley who authored one of the few works on the application of GTM in linguistics, laments this fact:

> In contrast to the spread of grounded theory in other fields of the applied social sciences, within applied linguistics, it has experienced marginalization and mistrust. (Hadley 2017: 4)

On the one hand, it has been argued that the assumption of ready-made categories poses an obstacle for sensible research in the domain of meaning (see 2.1.3). On the other hand, it is clear that no researcher can rid himself completely of theoretical assumptions before approaching the research area and some research experience is needed for any successful GTM study (Hadley 2017: 87). How can a balance between hypothesizing, which is the trigger for any research project, and the questioning of popular categories and their meaning be achieved?

In this respect, GTM can be valuable for contemporary variationist studies because the problematic assumption of a priori categories is handled explicitly. This notion is expressed as the avoidance of "forcing data" – in Glaser's words – and summarized by Charmaz:

> [F]orcing data includes: applying extant theories to the data; assuming the significance of demographic variables (such as age, sex, race, marital status and occupation; also called face-sheet variables) before beginning the study; and imposing evidentiary rules (a priori prescriptions about what stands as sufficient evidence) on the data. (Charmaz 2004: 62)

The relevance of this methodological principle for variationist studies which too often focused on "face-sheet variables" and especially for PVL is evident.

GTM stresses the use of methodological recursivity as a self-correcting process (Hadley 2017: 143): by the means of constant comparisons, it encourages the researcher to reflect his own position in the research process and to critically assess existing theories. Open exploration of the research area precedes a more focused investigation which leads to theory generation and in turn to further investigation, until theoretical saturation is reached. Different stages of research and theory generation are not only building on each other, but the constant comparison of data from different research stages sheds new light on existing data and its analysis. Decisions for the subsequent data collection are guided by recursive comparisons (Strübing 2008: 26 and Breuer 2010: 58). In GTM terminology this process of selective data collection is called 'theoretic sampling.' Essentially, there is no clear conceptual separation between data collection and analysis in GTM. Ideally, each piece of data needs to be analyzed, before moving on to the next event of data collection. Thereby, the criteria for the collection of data change as research proceeds. According to Hadley (2017: 41), "open sampling" should be followed by "relational and variational

sampling" and finally by "discriminate sampling". Charmaz describes the function of theoretical sampling for the data analysis as follows:

> Through theoretical sampling you can elaborate the meaning of your categories, discover variation within them and define gaps between categories. Theoretical sampling relies on comparative methods for discovering these gaps and finding ways to fill them. (Charmaz 2004: 78)

There are many possible methods which can be used for data collection in GTM. Commonly, qualitative methods are used, such as participant observation, the collection of media and print products for the purpose of communication analysis and interviews with experts and lay participants. More standardized methods and experiments can be used in advanced research stages. Therefore, quantitatively analyzable questionnaires, (perception) experiments and elicitation tasks, such as drawing mental maps or GERT, which is used in this study, can all be integrated in a GTM framework. Krefeld & Pustka's (Forthcoming) suggestion to use a triangulation of exploratory and more structured "hypothesis-testing elicitation methods" for the purpose of empirical studies on linguistic variation can therefore be achieved organically with GTM.

In qualitative research paradigms, 'coding' is used to designate interpretative analytic techniques (Strübing 2008: 19). In this sense, coding comprises rephrasing and summarizing textual data to explicate inherent concepts. In subsequent coding processes, these concepts are related to a larger amount of data and defined as analytic codes, if they are found to be appropriate for the specific context and type of data and relevant for the research questions. In GTM, there is no unique way of coding: because this interpretative technique is sensitive to the context and the type of data and is determined by the researcher's theoretic and personal dispositions, the coding process cannot be defined precisely and universally. Therefore, methodological works on GTM, such as Corbin & Strauss (2015), Hadley (2017) and Charmaz (2004), tend to explain coding in an illustrative fashion by the use of exemplary analyses. Although the approach qualitative content analysis, as described by Mayring (2015) and Kuckartz (2016), diverges from GTM in several aspects, both works were useful for this study because of their elaborate and detailed definition of qualitative analytic methods.

Generally, coding is understood in GTM as subsequent comparative processes which can be subdivided in "open" or "early," "axial" and "selective coding" (Strübing 2008: 20). Hadley describes the purpose of early coding:

> Coding at this stage serves not only to provide an accessible starting point for those new to the methodology of grounded theory but also simple, descriptive summaries of observable

> behavior and actions will be important for progressively building your theory. (Hadley 2017:
> 88)

An example for early coding in this study is the tentative formulation of categories on the basis of i53f2l1's statement (see 1). A more detailed account of the analytic processes which were used for coding the GERT corpus is given in 5.4. As a starting point for the analysis in this study, it was helpful to summarize the context and the Hebrew content of the interview events in English, in the form of case summaries, following Kuckartz's (2016: 58) method.

As can be seen from the examples which are provided by Charmaz (2004: 67), summarizing is an essential part of coding. Shortening and thus paraphrasing the content of interviews helps to get an overview and sets off interpretative processes which lead to the definition of analytic codes. At an early stage of analysis, it can be especially useful to look out for codes that have been brought up by the participants themselves. These are termed "in-vivo codes" in reference to "unique words or phrases used by [106] the informants[...] to encapsulate some important issue" (Hadley 2017: 105–106). Because of their authenticity which derives directly from the participants' formulations, they tend to persist throughout the analysis and can play an important role for theoretic reasoning.

'Axial coding' is characterized as a comparatively oriented way of analyzing data from a more encompassing perspective, with the aim to model relations between data and theoretic concepts (Strübing 2008: 20). The next higher level of analysis is reached through 'selective coding,' which is described as a readjustment of the analytical perspective:

> After selective coding, the analysis should be more consistent with regard to the research question than after axial coding.[22] (Strübing 2008: 22, my translation)

In GTM, the principle of recursivity which is implemented through constant comparisons is also present in the writing process. Before arriving at the final text, a systematic collection of memos is used to externalize one's thoughts during analysis. Birks et al. explain the term 'memo' as an acronym which refers to its functions:

> Mapping research activities; Extracting meaning from the data; Maintaining momentum; Opening communication [...] (Birks et al. 2008: 70)

Memos are not necessarily connected closely to the data, unlike codes, but can contain all sorts of methodological, analytic and theoretic reasoning (Hadley 2017: 107).

---

22 German original: Am Ende des selektiven Kodierens sollte aber die Analyse im Hinblick auf die Forschungsfrage ein höheres Maß an Konsistenz aufweisen als nach dem axialen Kodieren.

Memoing can be understood as a systematic method for entering into a discourse with the data and with oneself. By relating earlier and later memos to each other and in turn coding them – as the data itself, they can be used ideally for the conceptualization of a research paper. Writing memos can function as a trick to start the writing process early and without undue reservation because memos need not be orderly, stylistically elaborate or grammatically correct. They should just conserve the researchers' thoughts and associations at a certain point during analysis. Therefore, memos should at least contain a date, a topic and some text or a sketch.

In summary, GTM is a methodological attempt to explore RQs through the systematic integration of different perspectives on the data. In the context of this study, the overall RQ can be tackled by asking about the HSs' linguistic representations of themselves and their in-groups, in contrast to their representations of other groups. Once more, another perspective is added by the researcher who has to relate the insider and outsider perspectives by abstracting theoretic thoughts from the data. Ideally, these analytic processes of comparison and contextualization can lead to a "thick theorization" of the research area (Hadley 2017: 37). The overall aim of a GTM study is not to postulate an absolute and complete new theory, but to take in the reader as co-analyst, to propose interpretations and enable the "emergence" of theoretic thought in Glaser & Strauss's (1967) words (cf. Breuer 2010: 40). Hadley (2017: 11) points out that the style of writing used in GTM usually diverges from the typical academic style of the research domain. This less formal and more inviting style aims at creating an atmosphere of openness and enabling a more comprehensive understanding of the methodology and the research context.

The principles of GTM will be taken up again throughout Chapters 4 and 5, where I illustrate and discuss its possible application with several examples from this study.

2.3 Summary: Towards a Cognitive Variationist theory

The theoretic review encompassed works reaching from Ferdinand Saussure to Harvey Sacks and pursued the shifting focus from structure to speaker, leading to a usage-based study of meaning that is exemplified by third wave studies on linguistic variation.

On the basis of PVL, the argument has been made that any adequate study of linguistic variation needs to investigate the speakers' processes of social and linguistic categorization. To be able to determine the premises and the methods for detailed variationist studies, such as perception experiments, it is necessary to consider the institutions of the speech community and the speakers' common representations about their language and society. It has been argued that qualitative sociologist methodology and especially GTM are valuable resources for linguistic studies with

this aim. In GTM terminology, the researcher should look for 'in-vivo codes' and try to approximate the meaning of these categories, as they are constructed by the speakers.

Since processes of categorization are in the heartland of cognitive science, principles from cognitive linguistics such as prototypicality, markedness and basicness of level have much to offer for their study. Rosch's experiments on categorization were mainly carried out with concrete objects, but there is no objection against the validity of the principles for abstract objects (Schmid 2007: 125). In fact, units of a day, which are abstract concepts, were also studied as categories (Rosch 1978: 44). If concepts from our research context, such as 'dialect' and 'Mizrahim,' are socially constructed as "real" (cf. Berger & Luckmann 1967), the principles should be applicable, too.

This argument has already been made by Geeraerts (2008: 39) and Kristiansen (2008) who advocate for a paradigm of "Cognitive Sociolinguistics" and Krefeld & Pustka (Forthcoming) who suggest *A cognitive approach to language varieties* for the aims of PVL. Kristiansen (2008: 64) argues for the general treatment of "what in everyday terminology is referred to as language, accent, dialect, style and social group" as "concepts; categorizations and schemas on different levels of abstractions that relate to one another in the sense that they form part of a larger frame, or Cognitive Cultural Model."

This post-structuralist perspective on the popular notions 'language,' 'accent,' 'dialect' and 'social group' may seem counter-intuitive because they are conventionalized to such a degree – also (or especially) among linguists. However, differences in the treatment of these and related categories are well-known: in Israel, 'dialect' doesn't have the same importance in everyday life as in Europe and in the USA; the notion 'dialect' is used by non-expert speakers and linguists with another meaning than in European contexts (Sinner 2014: 16). In European studies about linguistic variation, the concept 'regionality' is much more prominent than in the USA and in Israel (in respect to Hebrew): Variationist studies tend to highlight the concept of regional varieties because they are most readily observable in a European context. How can these differing notions be integrated in a general theory about linguistic variation? Or should we give up on all the traditional notions?

It has been stressed that, at least, the speakers' understanding of these notions – better still, their own notions of systematic linguistic variation – need to be studied to gain valuable insights about their meaning and the dynamics of the system. Still, much work needs to be done to foster a theoretic framework which can fruitfully combine cognitive, sociological and linguistic strands of research. Since the most important common denominator of the research areas is their usage-based approach, much empirical research is needed – this should be understood as a call for a fourth wave of variationist studies (using Eckert's terms). This fourth wave could be profiled

as the study of speakers' representations of variation, which explicitly relates these concepts to cognitive science and the sociology of knowledge.

Ultimately, this approach can lead to a typological comparison of common (prototypical) concepts of linguistic variation in different speech communities, such as Levinson & Wilkins's (2006) collection of "grammars of space." In a global perspective, 'regionality' is not likely to be the universal defining variable for concepts of linguistic variation, as will be argued in the following analysis. Additional variables are at work which account for different LAs and language use among the speakers. These variables can be explored experimentally with context sensitive methods, such as GERT.

This study is an attempt to apply this theoretic approach with empirical data. The realignment of the research questions for this study along a post-structuralist perspective is illustrated in Fig. 2.3. Instead of departing from established sociological and linguistic categories, the entry point for this study are HSs' own processes of categorization. I argue that this approach is more promising and organic than a purely structuralist perspective because social and linguistic categorization are intertwined processes. To get a better understanding of the institutions which are
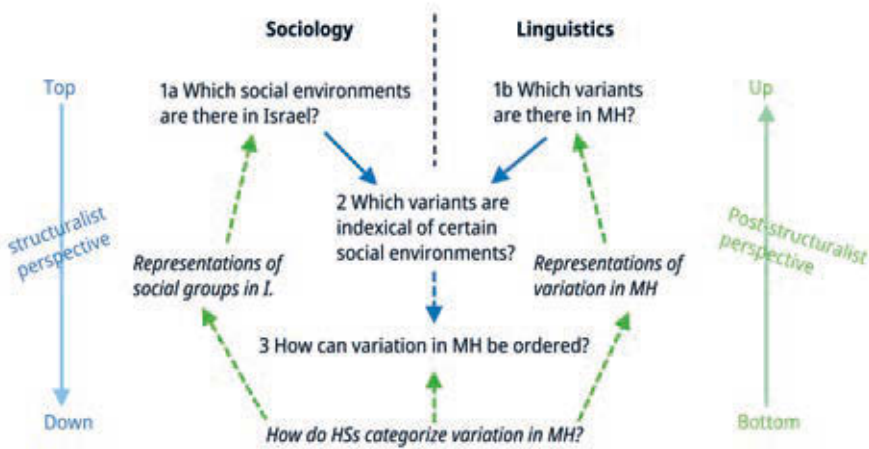

Fig. 2.3: Juxtaposition: theoretic perspectives on sociolinguistic variation

likely to affect HSs' representations and their language use, major historical and social developments in Israel will be reviewed in the next chapter.