Survey area selection in Variationist Linguistic Landscape Study (VaLLS)

A report from Vienna, Austria

Barbara Soukup
University of Vienna

This article addresses the unresolved issue of systematic survey area selection for large-scale quantitative Linguistic Landscape (LL) studies. It presents a strategy of 'hypothesis-driven stratified sampling' whereby survey areas are picked out in a nested, multi-step process on the basis of the configuration of local LL audiences (regarding age, multilingualism, and tourism) and ambient activity types (commercial vs. residential). The rationale for this strategy is drawn from variationist sociolinguistics; and the undertaking is accordingly cast as 'Variationist Linguistic Landscape Study (VaLLS)'. The details of the design are showcased and implications discussed in the context of the large-scale project 'ELLViA – English in the Linguistic Landscape of Vienna, Austria'. More generally, it is shown how the application of state-of-the-art variationist principles and methodology to quantitative LL research significantly enhances the latter's scientific rigor, which has been a major point of criticism.

Keywords: sociolinguistics, variationist sociolinguistics, linguistic landscapes, quantitative linguistic landscape study, data sampling, survey area, methodology, Austria, Vienna, English, German

1. Introduction

Wrapping up his comprehensive review of the state of the art of Linguistic Landscape (henceforth LL) research with an outlook on what's next for the field, Gorter (2013) identified three salient desiderata: (1) "to use empirical studies to test theoretical ideas rather than provide descriptive or analytic accounts that more or less illustrate theoretical ideas"; (2) to generally employ methodology that is "less explorative"; and (3) to conduct "rigorous research [...] that is well controlled and that can be replicated by other researchers" (Gorter, 2013: 205).

This call for greater methodological sophistication, systematizing, and rigor resonates particularly loudly in the context of quantitative LL research, a strand whose empirical practices, particularly of data collection, have been problematized virtually since its inception (see e.g. Amos & Soukup, 2020; Blackwood, 2015). Blackwood (2015) pinpoints two of the most contentious issues as the definition of the unit of analysis (the 'sign'), and the choice of a survey area (see also Backhaus, 2007). For the first, an arguably viable as well as widely practiced approach in quantitative LL research has been to follow Backhaus (2007: 66), and to take a sign to be "any piece of written text within a spatially definable frame" (see e.g. Amos, 2016; Blackwood, 2011; Edelman, 2014; Siricharoen, 2016; Soukup, 2016). While this is a working solution, the issue constitutes a pivotal methodological decision that is still ardently discussed in the field and far from settled. Thus, a detailed exploration would inevitably warrant a paper in its own right, which is why I suspend it here as beyond my present scope (but see Backhaus, 2007; Blackwood, 2015; Cenoz & Gorter, 2006, and Gorter, 2013 for critical reference; see furthermore Piritidis, 2014, for an illuminating study that directly tests for and identifies effects of different definitions of the unit of analysis on research outcome and findings).

For the second issue, the selection of a survey area, Blackwood (2015: 41) reports that there exists as yet "no obvious solution" that addresses the problem "in any definitive kind of way". At the same time, he points out that there is no way around it, because even in purportedly comprehensive, large-scale, quantitative LL studies, it is necessary to stake out a limited survey area for data collection: "It is challenging to the point of being unfeasible to survey an entire city or town" (Blackwood, 2015: 41).

This article sets out to redress the unsatisfactory status quo of the survey area issue in quantitative LL research. It introduces a rigorous strategy of survey area selection that is based on 'hypothesis-driven stratified sampling', as extrapolated from the rationale and toolkit of variationist sociolinguistics. In the following, I begin by discussing why and how variationist sociolinguistic theory and methodology provide a useful framework for consolidating quantitative LL surveys: for one, because they entail a tried-and-tested set of principles and strategies for sampling and data collection that, after decades of application and refinement, are highly sophisticated, systematic, and facilitative of replication (thus addressing Gorter's central desiderata listed above). I subsequently illustrate the application of such an approach in the study design of my own large-scale, quantitative survey of English language use in the LL of Vienna, Austria, under a project accordingly named 'ELLViA'. Specifically, I describe the study's hypothesis-based survey area stratification and selection scheme, and how a concrete set of expanses (street sections) were picked out for fieldwork within chosen areas. I then illustrate the

benefits of this methodological scheme with early findings, concluding with a round-up of my discussion and an outlook on its implications.

My proposed solutions to the survey area problem in quantitative LL research may not be equally 'obvious' (to use Blackwood's term) for all types of projects. Yet, I suggest that the presented scheme is flexible and general enough to allow for application across a wide range of settings and research interests, so that it may serve as a helpful blueprint for systematic quantitative LL investigations down the line.

2.   Integrating variationist sociolinguistics and quantitative LL study

On a foundational level, quantitative LL research shares a basic premise with variationist sociolinguistics: that insights into the "inevitable" (Fasold, 1990: 223) interaction between language use and social forces and dynamics can be gained from studying the distributional patterns of variants of expression across different social contexts.

Variationist sociolinguistics has traditionally focused on the investigation of spoken language (Schilling, 2013a), exploring how occurrences of particular variants of a 'linguistic variable' ("a set of alternative ways of 'saying the same thing'" – Labov, 1969: 738) systematically correlate with certain contextual parameters, such as, for example, speakers' regional background, age, gender, ethnicity, and/or sense of a situation or audience (see e.g. Chambers, 2008; Fasold, 1990; Guy, 1993; Kiesling, 2011; Labov, 1966; Milroy & Gordon, 2003; Schiffrin, 1994; Tagliamonte, 2006, 2012; Walker, 2010). Patterns of variant distribution are most commonly established by means of quantitative, statistical analysis of large numbers of tokens extracted from the speech production of certain speaker groups. Statistically confirmed patterns within the respective variable matrix are regarded as both scientifically and socially meaningful. They serve as the basis for detailed descriptive and, ultimately, explanatory statements about the nature of the relationship between linguistic and social structure.[1]

In a similar vein, quantitative LL research has been conceptualized as the study of "the distribution of [LL] items, uses of languages, categories of designs and texts that unveil the relative impact of different structuration principles" (Ben-Rafael, Shohamy & Barni, 2010: 344). In other words, while the focus here

---

[1]. This, of course, refers to Weinreich, Labov, and Herzog's (1968) seminal notion of the 'orderly heterogeneity' of variable language use, meaning the fact that there is a certain identifiable order and system to variation and its constraints, and not mere unpredictable chaos. See e.g. Kiesling (2011) and Schiffrin (1994) for further discussion.

is typically on written language use ("any display of visible written language" – Gorter, 2013:190), insights are, just as in variationist sociolinguistics, generated by examining and interpreting constellations and patternings of linguistic variables (notably, language choice and forms of multilingualism) relative to social variables (authorship, audience, sign appearance, placement, thematic context – see e.g. Amos & Soukup, 2020; Backhaus, 2007; Blackwood & Tufi, 2015; Gorter, 2006).

On the basis of this shared premise, quantitative LL research can be conceptualized as a variationist sociolinguistic enterprise – so that it can be cast as 'Variationist Linguistic Landscape Study' ('VaLLS' for short – see Amos & Soukup, 2020; Soukup, 2016). And on these grounds, tried-and-tested variationist theory and methods can be harnessed and made fruitful for the purposes of quantitatively oriented studies of LLs.

One particular affordance of this transfer is that variationist sociolinguistic methodology features standard precepts of data collection that have been shown to ensure research validity and replicability (two yardsticks of scientific rigor and quality in quantitative methodology; see also Dörnyei, 2007). Most important here is the 'Principle of Accountability' (Labov, 1969). Its purpose is to put the occurrence of linguistic variants in a sample into proper perspective, and to prevent over- or understating the salience of a variant by way of selective focus or anecdotal reporting (e.g. due to the allure of the variants' markedness, exceptionality, or non-standardness; see Labov, 1969:737–738). Thus, the Principle of Accountability holds that "any variable form (a member of a set of alternative ways of 'saying the same thing') should be reported with the proportion of cases in which the form *did* occur in the relevant environment, compared to the total number of cases in which it *might have* occurred" (Labov, 1969:738; my formatting). In the words of Tagliamonte (2006:13), "you cannot simply study the variant forms that are new, interesting, unusual or non-standard [...]. You must also study the forms with which such features vary in all the contexts in which either of them would have been possible." It is the application of this principle that allows for a scientifically robust generation of findings predicated on statistical comparison of the rates of occurrence of particular variants across different contexts (i.e., analysis of frequency distributions).[2] Such comparison must use a standardized and normalized measure of the frequencies at which different variants occur in

---

2. Note here the fundamental tenet of variationist sociolinguistics that distributions of linguistic variants across different contexts are rarely truly complementary, but almost always probabilistic: it is the *likelihood* of one variant occurring over another that changes across contexts. This is because heterogeneity (rather than categoricity) is seen as an intrinsic feature of all language use, albeit a 'structured' feature that can be investigated for its systematicity (see again

particular data subsets, rather than absolute counts stemming from what are usually disparate reference pools. This measure is typically the percentage of occurrences of a particular variant within the entire set of occurring variants of the same linguistic variable in a given context (i.e. within all recorded alternative ways of 'saying the same thing'). If it is then confirmed that one variant is more likely to occur (i.e. occurs at a higher normalized rate/percentage) in one particular type of context than in another, this is considered evidence in favor of the prediction of the underlying test hypothesis (rejecting the null hypothesis that the distribution is the same), and an empirical basis for interpreting the meaning of this association, in terms of how it may reflect and/or construe broader social dynamics. Kiesling (2011) provides a list of 'canonical' findings of sociolinguistic patterns that have thus been established over the past decades of variationist research.

In short, the Principle of Accountability is the bedrock on which the scientific validity and rigor of quantitative variationist research protocols rest. It enables statistical hypothesis-testing regarding distributions of normalized frequencies of variants across different social contexts (e.g. different speaker groups). Its precondition is a count-all procedure of data collection across a sample, meaning that all possible realizations of a linguistic variable must be encompassed.

For VaLLS projects, the linguistic variable under investigation is typically language choice, and its variants are local minority and majority languages – for example, whether a text is written in Hebrew, Arabic, or English in Israel (Ben-Rafael, Shohamy, Amara, & Trumper-Hecht, 2006), Basque or Spanish in the Basque Country, and Frisian or Dutch in Friesland (Cenoz & Gorter, 2006), Japanese or a language 'foreign' to the country in Japan (Backhaus, 2007), French or the respective regional heritage language in Brittany and Corsica (Blackwood, 2011), Chinese or (mainly) English in a UK Chinatown (Amos, 2016). Written signs by definition choose and use some kind of a textual system of linguistic expression (language, variety, alphabet). Thus, data collection entails recording all visible signs in a survey area that bear any text on them.[3]

Just as it is next to impossible to survey entire cities in LL research, it is quite unrealistic in most variationist sociolinguistic projects to exhaustively survey the

---

Weinreich et al.'s 1968 notion of 'orderly heterogeneity', and its discussion in, i.a., Chambers, 2008; Kiesling, 2011; Tagliamonte, 2006, 2012; Walker, 2010).

3. This point once more throws into relief the centrality of the definition of a 'sign' as a unit of analysis for quantitative LL research (see my point further above). It underscores the fact that, in order to avoid researcher bias and circularity, this definition needs to be laid down prior to data collection, in objective terms independent of findings, rather than in any ad hoc terms (relating e.g. to sign salience) that jeopardize the validity of statistical analysis. This is what makes (a strict application of) Backhaus' (2007) approach particularly viable for VaLLS (see also Soukup, 2016).

entire population one is interested in (Schilling, 2013a). For this reason, sampling (picking out particular sets of speakers for detailed study) plays a very important role in variationist sociolinguistics, too; and careful decision-making is essential. In other quantitative sociologically-minded paradigms (e.g. demographic or economic research), strict random sampling, or the selection of a set of informants that are truly representative of a population at large, is the gold standard, findings being of relevance only in so far as they may be generalized from the sample to the entirety of the population it is taken from. However, in variationist studies, the focus is on "detecting and measuring differences between different *subgroups* of the population" in terms of their language use (Sankoff, 2005:1001; my emphasis). As Sankoff explains, from a variationist perspective,

> Indeed, we often prefer to regard as an empirical problem the determination of the social dimensions along which linguistic change and variation proceeds. In this case the appropriate strategy is to ensure that as much as possible of the existing linguistic diversity in the community is represented in the sample. This is done by assuring that whatever auxiliary variables we suspect may be correlated with some aspect or other of linguistic variation, such as age, sex, place of birth, etc., are represented as fully as possible in the sample. To accomplish this, a stratified design is set up prior to sampling. The idea is to divide the population into a number of strata, each of which contains only individuals falling into a restricted range on one or more of the auxiliary variables. Thus, one stratum might contain all women of a certain age range born in a certain district. […] In the most elementary sense, the sample is not representative, but from a more sophisticated point of view, the stratified sample is more informative than a completely random one would be. (Sankoff, 2005:1001)

In other words, sampling in variationist sociolinguistics is typically based on a judicious process that operationalizes expectations (hypotheses) about the interaction between language use and social factors in the form of selection criteria for informant groups to be surveyed. The ulterior methodological rationale is therefore one of hypothesis-testing across a pre-established independent variable matrix, regarding distributional patterns of a set of linguistic variants (i.e., the dependent variable). Application of this technique in lieu of true random sampling has been further justified by the finding that "linguistic usage is typically more homogeneous than other behaviors," so that "we can usually be confident that the insights we obtain [via stratified sampling] are generalizable to the larger population" anyway (Schilling, 2013a: 33/35). In addition, due to its very nature "there is no guarantee that a strictly random sample will yield data from speakers in all the categories of interest, particularly if certain segments of the population are statistically underrepresented (e.g. ethnic minorities, residents of remote communities)" (Schilling, 2013a: 33).

My proposed solution for the survey area problem in quantitative LL research is to adopt and adapt also in VaLLS surveys the outlined technique of hypothesis-driven stratified sampling, as commonly used in variationist sociolinguistics (whereby, to recap, informants are selected from hypothesis-based groupings within a population). This strategy in fact addresses all three of Gorter's initially listed desiderata ("to use empirical studies to test theoretical ideas," to employ "less explorative" methodology, and to conduct "rigorous research" – Gorter, 2013: 205). At the same time, it is a way to keep a large-scale VaLLS project manageable while implementing the requisite count-all procedure of data collection, which may be quite demanding regarding human resources and fieldwork hours.

In Section 3 below, I provide a concrete example of how to put the technique into practice in a workable and analytically functional fashion, as I describe the study design of the VaLLS project 'ELLViA – English in the linguistic landscape of Vienna, Austria'.

3.   Hypothesis-driven stratified sampling in ELLViA

In line with much preceding quantitative LL research (e.g. Amos, 2016; Backhaus, 2007; Ben-Rafael et al., 2006; Blackwood, 2011; Cenoz & Gorter, 2006; Huebner, 2006), the survey area for the ELLViA project is composed of stretches of streets. Along these, all visible manifestations of written language were to be recorded. Therefore, it was the selection of these stretches that constituted the sampling issue to be resolved. This involved two basic steps: (1) hypothesis-based stratification of Viennese sub-areas in which to select streets, and of streets to be selected; and (2) picking out concrete stretches on these streets, in an *ex ante* fashion (i.e. conceptually prior to and independent of data collection). Each of these steps is now described in turn.[4]

3.1   Hypothesis-driven stratification and selection of sub-areas and streets for surveying

As its name suggests, the general goal of the ELLViA project is to investigate the manifestation and communicative function of English language use in the LL of Vienna, Austria (for details on the full project see Soukup, 2016). The project consists of three modules, two of which investigate folk perceptions of English in Vienna. My concern here is only with the third one (first in chronological

---

4. For additional details like source data, tables, and exact figures, omitted here for reasons of scope, contact the author.

order), which consists of a comprehensive VaLLS survey of written signs that can be found in the Viennese LL, in view of statistical extraction of patterns regarding English versus German language use (= the dependent variable). Research hypotheses for fieldwork and data collection thus express the expected impact of certain pre-determined factors on the presence or absence of English on written signs in Vienna.

A pivotal assumption underlying hypothesis-generation in ELLViA is that local audiences of potential sign-readers (= expected passers-by) bear a direct influence on language choice on LL signage. This premise is derived from the rationale of an interactional, dialogical model of communication (see *inter alia* Bakhtin, 1986[1952–53]; Erickson, 1986; Goffman, 1959; Gumperz, 1982, 2001; Linell, 2017; Soukup, 2016; Tannen, 1989, 2004). Under such a model, any kind of communication is conceptualized as a meaning-making process based on mutual anticipation, interpretation, and negotiation between addressers and addressees. Both play an equally important role as active participants who jointly create and make sense of what is going on. Their relationship is 'dialogical' in that it is of a two-way nature: where addressers design their expressions in expectation of addressees' responses, trying to influence these responses (i.e. trying to give rise to certain communicative messages), addressees in turn are not merely passively subjected to addressers' expressions but actively shape them through their responsive stance. Arguably, this model also applies to the production and reception of written text in the LL, which can thus be regarded as the physical manifestation of an (albeit rarely face-to-face) interactional dialog between sign authors and sign readers.[5]

One important consideration of sign authors within this dialog is that language choice in and of itself gives rise to certain messages, as it 'contextualizes' (Gumperz, 1982) a text with the social meanings (language attitudes, ideologies) associated with the chosen language. Thus, sign readers can be expected to activate the social meanings of a particular language choice in the course of their interpretive process of making sense of LL sign communication. Strategic, deliberate language choice on signs is therefore a way to push certain communicative messages, trying to both anticipate and influence reader response. Soukup (2009) is a study of how these same processes play out in spoken language use and language choice (code-switching) in the Austrian context.

It is the assumption of said inherent dialogicality of sign generation and reception that paves the way for research hypotheses which relate language choice

---

5. See in particular Linell (2017) for discussion of this conceptualization of 'dialogicality' and its applicability to written language. For a general review of literature on the inherent dialogicality of reading written language, see Reichl (2009).

on LL signage to the make-up of a population local to the emplacement of a sign. In other words, such hypotheses operationalize and empirically test the idea that an LL sign is to a considerable degree linguistically geared towards the people most likely to pass by and read it (see also Spolsky & Cooper, 1991). Establishing hypotheses about the use of English on LL signage in Vienna therefore also involves establishing what kind of local audiences are most likely targeted and reached by such use.

Relevant literature suggests that English language use in European 'foreign-language' settings can be associated with at least two social parameters: (1) age – in the sense that English shows a higher presence in the communication of young people versus older people (e.g. Berns, de Bot, & Hasebrink, 2007; Grau, 2009; Hilgendorf, 2007; Hofmann, 2002; Leppänen et al., 2011; Pitkänen-Huhta & Nikula, 2013); and (2) multilingualism – in the sense that English frequently serves as a 'lingua franca' in settings and situations where people of different linguistic backgrounds come together (e.g. Seidlhofer, 2011) – including tourism activities (see e.g. Bruyèl-Olmeido & Juan-Garau, 2009). In addition, English nowadays bears a strong association with international and global consumerism and commerce, particularly as regards its use in advertizing (see e.g. Backhaus, 2007; Kelly-Holmes, 2000, 2005; Piller, 2003). This set of considerations was molded into the research hypotheses underlying the ELLViA sampling scheme. Accordingly, it was hypothesized that *more* English would be found on Viennese LL signs recorded in (1) survey areas that are inhabited by a local audience with a high proportion of young adults, versus areas with a high proportion of senior citizens; (2) survey areas with a linguistically highly diverse population, versus a linguistically fairly homogeneous (German-speaking) one; (3) survey areas with high amounts of tourist foot traffic, versus ones with rather low amounts; and (4) survey areas where the local population's role as consumers is foregrounded – i.e., commercial survey areas, versus predominantly residential ones.

Selection of the final set of survey areas was based on a nested decision-making process that took Vienna's official, twenty-three administrative districts ('Gemeindebezirke') as its starting point (see the official listing of the City of Vienna: https://www.wien.gv.at/bezirke/ – retrieved October 20, 2019). It was carried out from late 2014 to early 2015, using the most recent demographic data available at the time, mostly from the Austrian national statistics institute Statistik Austria (responsible for census-taking), and publications by the city government of Vienna (notably the *Statistisches Jahrbuch der Stadt Wien – 2014* – Magistrat Wien, 2014).

For the variable AGE (of the local population), all administrative districts were ranked according to their proportion of young adult residents (operationalized as 20–30 year-olds, i.e. 'digital natives' [cf. Prensky, 2001] with already some auton-

omy of action and economic clout), as well as according to their proportion of senior citizen residents (operationalized as 65+ year-olds; 65 being the maximum legal retirement age in Austria), and according to the average age of the local population (to control other bias).[6]

For the variable MULTILINGUALISM, proxies of residents' citizenship and native origin had to be used. This was because (home) language background was not recorded in the most recent census (2011), nor was the information publicly available elsewhere; and the 2001 census data, which featured a respective question, were deemed obsolete, due to a large influx of immigrants in the interceding decade: between 2001 and 2011, the percentage of Viennese residents with a non-Austrian place of birth increased from 24% to 31% (Statistik Austria, 2013: 51). Administrative districts were thus ranked according to their percentage of residents with non-Austrian citizenship, as well as of residents born in a country other than Austria.[7] Additional pivotal input and perspective were provided by Katharina Brizic, who had conducted a comprehensive survey of (self-reported) language use among Viennese elementary school students (3rd and 4th grade) in 2009, covering 85% of all Viennese elementary schools (see Brizic, 2013; Brizic & Hufnagl, 2011). She kindly and generously shared relevant (anonymized) data with me (Brizic, 2015, p.c.) that allowed for a rough (yet the best available) calculation of the percentage of multilingual children per administrative district. Districts were also ranked according to this statistic.

For the variable TOURISM, the proxies of registered visitors (i.e. persons who spend at least one night) per district resident, as well as each district's share of all visitors to Vienna, and each district's recorded number of museum visits were used, and districts ranked accordingly. The last variable, COMMERCIAL ACTIVITY,

---

6. In Austria, the current legal retirement age for women is 60 and for men 65 (though the actual average retirement age in 2015 was 61.3 for men and 59.2 for women – see Statistik Austria 2019). The higher age of 65 was picked as a benchmark here because practically *all* residents at and beyond that age could be expected to be retirees.

7. Of course, some residents with non-Austrian place of birth and citizenship may still be native speakers of German – most notably, those originating from Germany and (parts of) Switzerland. Data available for residents' country of birth (from Stadt Wien – data.wien.gv.at: https://www.wien.gv.at/statistik/ogd/vie_107.csv, retrieved on January 3, 2017) showed that in 2014, 2.6% of Viennese residents were of German origin (data for Switzerland not available). This seems to be a negligible rate overall; and the factor was furthermore controlled in the sampling scheme for the variable MULTILINGUALISM by matching districts with similar, low ratios of German-born residents. Incidentally, the rate of non-Austrian residents who may be native speakers of English is even more negligible, with only 0.5% nationals stemming from 'inner circle' (Kachru, 1992) English countries (UK, USA, Canada) who, additionally, do not live in any particular enclaves.

was operationalized via the street selection itself within each chosen district (a point I return to below).

The selection process for the set of districts to be featured in the sample was subsequently based on picking out those pairs where one district loaded particularly high and one particularly low on one of the three above-mentioned groups of factors (AGE, MULTILINGUALISM, and TOURISM), but within the Viennese average on the other two. In each pair, the two districts' values on the relevant test variable (i.e. respective percentages of the population) were apart by at least one standard deviation from the mean.

Further considerations led to the exclusion from analysis of the 9th district (because of its high density of tertiary education institutions, deemed a potential confounding factor regarding English language usage), as well as the 13th district (for loading both very high on 65+ year-olds and very low on MULTILINGUALISM), the 22nd district (because the United Nations Viennese headquarters are located there as well as some of its international employees' residences, also considered a potential confound), and the 23rd district (a very large district known for the general heterogeneity of its residential population, as confirmed by Viennese urban sociologist Christoph Reinprecht, 2015, p.c.). The 1st district also loaded very high on two variables simultaneously (65+ year-olds and tourism); however, because it is the single biggest tourist destination in Vienna by far, it was retained in the set.

At the end of the selection process, the final set comprised (1) for AGE, the pair of the 8th district (representing the young adult population) and the 19th district (representing the senior citizen population); (2) for MULTILINGUALISM, the 16th district (representing a multilingual population) and the 21st district (representing a fairly monolingual German one); and (3) for TOURISM, the 1st district (representing high tourist activity) and the 18th district (representing low tourist activity). Care was taken that the districts were paired not only in terms of the respective test variable, but also that they match regarding the socioeconomic characteristics of the resident population (as again confirmed by Reinprecht, p.c., 2015; see also Rode, Schier, Giffinger, & Reinprecht, 2010), which was thus controlled as a potential confound.

The next step in the nested sampling process was to operationalize the fourth variable, COMMERCIAL ACTIVITY, relating to the hypothesis that English would show an increased presence in commercial versus residential LLs. Thus, in each selected district, one predominantly commercial and one predominantly residential street were picked out. To identify those streets with highest commercial activity, the ratio of commercial establishments per meter was calculated for each street in a given district, using data kindly provided free of charge by Herold Business Data GmbH, in the form of the number of yellow-pages (i.e., registered business) entries for each street. Street lengths were compiled from measurements provided

by the city government of Vienna via their open data initiative.[8] Because, as a result of pilot fieldwork, the length of the stretch of street to be surveyed was set at 200m (squaring considerations of representativeness and feasibility), only streets of a minimum length of 250m were included in the calculations.

Thus, a total of six streets with the most businesses per meter – one street for each selected district – were chosen to represent high commercial activity and consumer-orientation in the sample.[9] For the residential streets, the focus was laid on those streets with the lowest number of businesses per meter, but demographic stratification was additionally fine-tuned using census sub-districts ('Zählbezirke') as a basis. This means that in each administrative district of the overall set of six, those sub-districts were identified that loaded particularly high on the variable the entire district was chosen to represent, so that the local population could be pinned down on an even smaller level of granularity. At the same time, a selected residential street needed to be as close as possible to its counterpart, the shopping street, to ensure a shared LL audience. A further restriction on residential street selection was that they should match the shopping streets as far as possible regarding the type of urban architecture. This meant that only those residential streets were taken into consideration that, like the shopping streets, feature a fairly closed stretch of mural façade, rather than lots of single-family homes set far back from the street, with front yards and behind fences. To ensure that the chosen settings would fulfill all required criteria, street selection

---

8. Street length data retrieved online from https://www.data.gv.at/katalog/dataset/1039ed7e-97fb-435f-b6cc-f6a105ba5e09 (April 10, 2015). No official or public data on businesses per street were available that matched the quality and comprehensiveness of the Herold data.

9. N.B.: Using shopping streets within a demographically stratified sample of districts as a gateway to the commercial (shopping) activities of local residents in Vienna is further supported by Viennese city governmental policies that explicitly (financially) support 'Nahversorgung', or the availability of basic shopping amenities for residents close to where they live (see e.g. Wirtschaftsagentur Wien 2017).

By name, the final set of shopping streets featured in ELLViA are: Graben (1st district), Josef-städter Straße (8th district), Thaliastraße (16th district), Währinger Straße (18th district), Döblinger Hauptstraße (19th district), and Am Spitz (21st district). The last is in fact a triangular square of sorts, whose outer perimeter was surveyed (in its center sits a single, huge building housing the district administration), yielding a single street-side length of about 400m (to match the double-sided 200m of the other stretches). Döblinger Hauptstraße and Am Spitz were actually those streets with the second highest businesses-per-meter ratio in their district. However, in the 19th district, the first street in rank was rejected because its high ratio was found to be caused by one single office tower that had a great number of individual businesses registered. In the 21st district, the first street in rank is a huge open square-cum-public transportation hub whose surveying was not deemed reasonably feasible nor its make-up comparable to that of the other streets in the sample.

'on paper' was complemented by extensive site visits. This way, another set of six street survey areas – this time with very little commercial activity – were picked out.[10]

The entire ELLViA survey area thus encompasses a total of twelve street sections, two in each selected district. A geographic overview is provided in Figure 1.
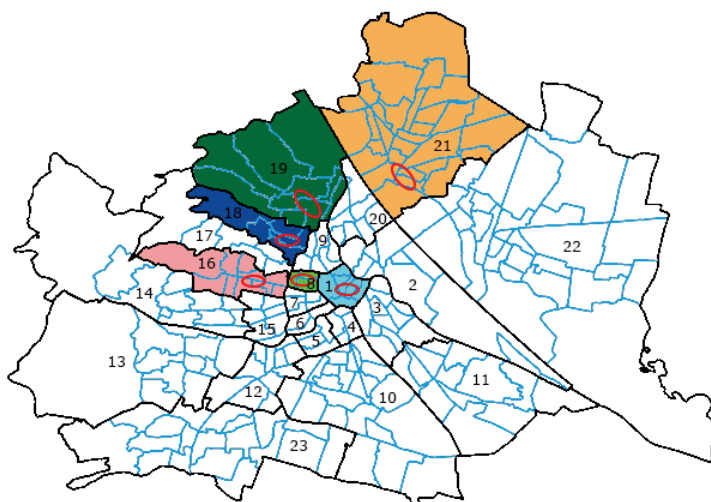

Figure 1. Map of Vienna featuring administrative and census district borders and showing the approximate location of ELLViA survey areas (map adapted from Stadt Wien – data.wien.gv.at, Katalog Zählbezirksgrenzen Wien)[11]

---

10. Note that in the 1st district, the Inner City, it turned out to be impossible to find a street of adequate length (≥250m) that does not feature a commercial establishment in every single building. Likewise, the 19th district features lots of single family homes with surrounding gardens. Thus, in both districts, the survey area was eventually patched together by incorporating stretches from more than one shorter residential street within the relevant census district, which was seen as preferable over not upholding the other basic selection criteria.

The final set of residential streets comprises, in the 1st district, the area of Domgasse-Blutgasse-Grünangergasse-Kumpfgasse (close to St. Stephen's cathedral); Stolzenthalergasse in the 8th district; Thalhaimergasse in the 16th district; Plenergasse in the 18th district; Iglaseegasse and Pfarrwiesengasse in the 19th district; and Kinzerplatz (a broad square whose outer perimeter was surveyed and which runs around a single huge church in the middle) in the 21st district.

11. https://www.data.gv.at/katalog/dataset/e4079286-310c-435a-af2d-64604ba9ade5 (retrieved on December 29, 2017)

As already mentioned, pilot research suggested that a two-sided street section of 200m (or 400m one-sided single-line total) makes for a data collection area that is meaningful and informative while also being manageable in terms of field-work. Section 3.2. now describes the procedure of picking out such stretches of street for the sample, as the third step in my nested survey area selection process.

3.2   Picking out stretches on the streets to be surveyed

The main challenge in picking out concrete, clearly delimited stretches on the selected streets to be surveyed was to find a method that could be applied to all streets in the exact same way, and in an *ex ante* fashion (i.e. conceptually prior to and independent of data collection) that would retain an element of randomness and avoid researcher bias (e.g. towards sign density, salience, or make-up). The solution adopted was to identify the lengthwise mid-point of each street, and then to stake out 100m up and down the street from this point along the street axis. This approach can be described and theorized as 'systematic cluster sampling'. It is 'systematic' because, unlike strict random sampling, data collection proceeds on the basis of a pre-defined systematic parameter – here, selecting a street's mid-point as the center of measurement and point of departure for sampling. It is a process of 'cluster sampling' because it uses units consisting not of individual, randomly spaced data points (LL signs), but of groups of them in adjacency, in the form of all signs along a continuous stretch of street (see e.g. Levy & Lemeshow, 2008, and Thompson, 2012 for reference).

Measurements to stake off the ELLViA survey areas were carried out using the interactive city map of Vienna that is provided online, free of charge, by the city government, and conveniently features a virtual measuring tool (https://www.wien.gv.at/stadtplan/ – last accessed January 4th, 2018). Figure 2 is a consolidated illustration of the steps taken in the delimitation process for the residential street in the −AGE district (Stolzenthalergasse). As a first step (labeled [1] in Figure 2), the entire street was measured lengthwise along the street's mid-axis (as corresponds with the official measurement protocol of the city of Vienna's governmental city surveying – MA 21, p.c., 2015). The street's mid-point was identified by simply dividing the entire length in half and reapplying this measure to the map. Secondly ([2]), the 200m survey area was roughly delimited by measuring 100m up and down this midpoint. Finally, to set precise, clear, and replicable beginning and end points for data collection, the survey area boundaries were set at edges of entire buildings close to the 100m pegs up and down a midpoint. A measurement of the actual survey area length was then taken along the lengthwise street axis, for each street side separately (step [3] in Figure 2). Measurements were double-

checked via pdf-printouts of the map that render the 1:1000 scale more precisely than the interactive online version.
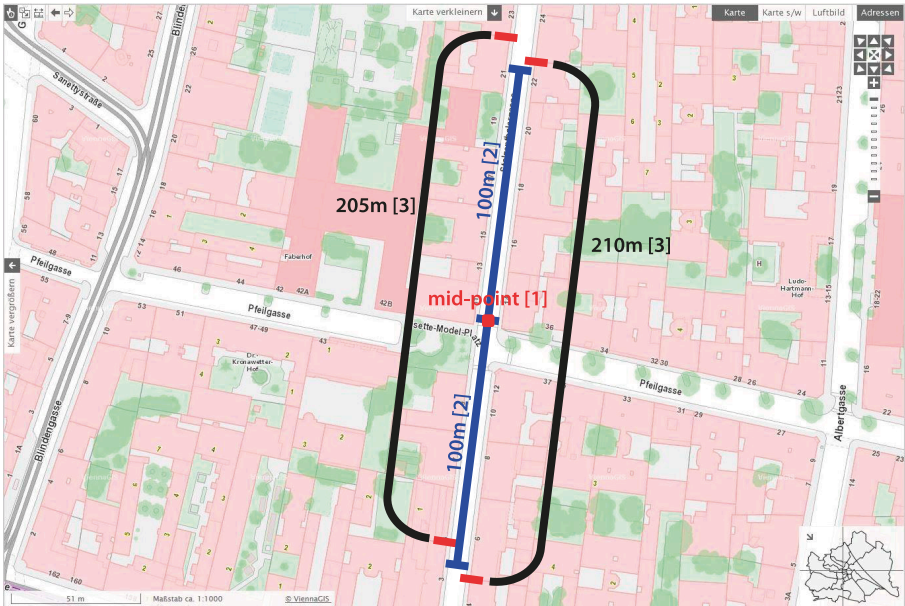

Figure 2. Steps [1]–[3] of measurements and boundary setting for the survey area around the midpoint on Stolzenthalergasse, as implemented in ELLViA (screenshot; map source: Stadt Wien – ViennaGIS, www.wien.gv.at/viennagis/ – last accessed January 17, 2018)

Table 1 displays the final set of ELLViA survey areas in terms of their length (per one single street side), for each test variable operationalized via the different districts, as well as for the two types of street.

Table 1. ELLViA survey area (street) section lengths (single-sided, rounded to full meters) per independent variable category

	+AGE (=old)	−AGE (=young)	+MULTI LINGUALISM	−MULTI LINGUALISM	+TOURISM	−TOURISM	TOTAL
+COMMERCIAL ACTIVITY	396m	407m	409m	401m	400m	395m	**2,408m**
−COMMERCIAL ACTIVITY (residential)	402m	415m	410m	388m	442m	397m	**2,454m**
TOTAL	**798m**	**822m**	**819m**	**789m**	**842m**	**792m**	**4,862m**

The ELLViA survey area thus comprises a total of 4,862m of street (single-sided). Some further considerations beyond the already listed fed into setting up this total, as a function of the need to deal with contingencies in situ while upholding the comparability of the selected street sections across the sample. For one, in the 16th and 18th districts, sampling skipped stretches of one large square-cum-city-park each (which 'interrupts' the shopping streets there, also in terms of street numbering). Measurement of the respective 100m stretch to each side of the midpoint leaves off before and picks up again after. Further, at any intersection, data collection did not venture into crossing streets at any depth but kept to a projected, virtual, straight line between the façade corners on the surveyed street. This rule was suspended only for buildings with canted corners: for these, the corner façade set at an obtuse angle to (and thus not fully in alignment with) the façade fronting the surveyed street was still included in data collection, if it was part of a corner shop with shop windows onto the surveyed street.

As already mentioned, data collection during fieldwork proceeded along the outer façades of all buildings in the selected street sections. Entrance ways of buildings were covered up to the first outer barrier (e.g. door or gate). For each street side, data collection extended to the middle of the lane (thus including, for example, manhole covers located in the roadway).[12]

All signs that were visible from street level and, as far as could be determined, at least half complete (e.g. when torn off) were included in data collection, i.e., recorded via photography and paper-based field-note tables, including assignment of a unique identifier. No limitations of placement height were imposed (see also Backhaus, 2007). For those signs that visibly featured some text that was, however, not clearly legible to the naked eye, a camera zoom was used to record textual details (under the argument that such technology would nowadays be readily available to any layperson taking a close interest in the LL).

Fieldwork was carried out in a team of two (one photographer, one field-note taker) from April through September 2015, on workdays during business hours

---

12. Photography of the outer façades is in line with legal privacy rules in Austria that define the outsides of buildings as open to public photography (§ 54 Absatz 1 Urheberrechtsgesetz). To ensure full compliance with privacy regulations, no photos were taken of doorbell panels and apartment mailboxes with private name tags. In addition, four ATMs (cash dispensers) were omitted due to security concerns of respective local bank management.

Note that manhole covers and other items on the streets were recorded only insofar as they were visible at the time of fieldwork, meaning that any item invisible due to cars parking on top are not included. No remedial field visits were carried out in this regard, as, at any rate, LL research can only ever capture an ad hoc snapshot of ever-changing surroundings (stickers may be covered with other stickers, ad posters may be exchanged etc.). This, of course, constitutes a known limitation of the count-all procedure (or indeed most data collection procedures) in LLs.

(9:00–18:00). It took up a net total of 424 woman-hours (212h per person) – 330h for the shopping streets and 94h for the residential streets. In our team of two, fieldwork thus consumed on average 2–3 minutes per meter covered (a little over 4 minutes for the shopping streets, and a little over 1 minute for the residential streets). Photographer and note-taker initially ventured forth separately on the same street to record items but came back together at intervals to match up their findings, thus also ensuring an immediate, on-site, four-eyes check on data collection for reliability and completeness.

The camera used was a Panasonic Lumix DMC-TZ61 digital with 30x optical zoom. Pictures were recorded as JPEGs at the camera's 'finest' quality setting (18 megapixels/ 4896×3672 pixels). During fieldwork, both researchers wore academically branded T-shirts and nametags to openly signal the type of activity going on (i.e., research) to passers-by and shop staff. The latter were also personally informed of our undertaking prior to photography, to prevent suspicion and secure cooperation where necessary (e.g. closing open doors for a better picture). A clipboard was used for field-note taking and additional signaling of purpose, and business cards of the PI offered for follow-up contact and questions.

Full data coding and analysis of the final, comprehensive ELLViA dataset were still ongoing at the time of writing this article. In the following section, I briefly present some basic frequency distributions that are already available, to showcase the most salient benefits of the sampling scheme implemented and described here.

4.   The benefits of the sampling scheme: First results from the ELLViA VaLLS survey

In the course of fieldwork in the established survey areas, a total of 17,109 LL items ('signs', 'things with written language on them') were recorded. These now constitute the full ELLViA dataset. An immediate benefit of the precise circumscription and measurement of the survey areas, as well as the count-all data collection procedure, is that this permits a general 'LL density' calculation: overall, an average of 3.5 LL items were recorded per street meter covered in Vienna. Future studies in different cities could use this measure as a point of comparison across settings.

The ability to normalize item counts per street meter furthermore paves the way for quantitative, distributional comparisons across different subsamples of the dataset on an equal footing. In the case of ELLViA, this notably involves comparisons across the different Viennese administrative districts and street types operationalizing the main research hypotheses (see Section 3.1.). Table 2 presents the LL densities in the respective survey sub-areas.

Table 2. LL item densities in the ELLViA sub-survey areas (labeled for the relevant independent variable category), ranked from highest to lowest total per district

	+TOURISM	−MULTI LINGUALISM	−TOURISM	+AGE (=old)	−AGE (=young)	+MULTI LINGUALISM	TOTAL
+COMMERCIAL ACTIVITY	9.3	7.3	5.2	4.7	4.2	4.2	**5.8**
−COMMERCIAL ACTIVITY (residential)	1.5	1.7	0.7	1.2	1.7	0.8	**1.3**
TOTAL	**5.2**	**4.6**	**3.0**	**2.9**	**2.9**	**2.5**	**3.5**

As Table 2 summarily shows, we find that the set of the commercial streets yields a higher average, of 5.8 items per meter, than the residential streets, with only 1.3 items per meter. The same pattern is also manifest within each individual district. At first glance, it may seem unsurprising and intuitive that there should be more LL items present in a commercial than in a residential setting, on the expectation of heightened promotional signaling activity in the former. Yet the results provided here are noteworthy in that they represent not merely impressionistic but rather robust comparative evidence, given that they are based on frequency counts normalized across the different locales, and draw on a rigorous, systematic, and exhaustive survey on an unprecedented level of granularity (down to, for example, tiny bolts that are used to fix trash cans to traffic poles, and that also bear some text).[13]

Table 2 furthermore shows that the commercial street in the +TOURISM (the 1st) district exhibits the highest LL item density, at 9.3 items per meter, followed by the −MULTILINGUALISM (the 21st) district with 7.3 items/m. In the case of the former, the featured commercial street ('Graben') is in fact a broad pedestrian zone with many outdoor café areas, and also one of the most highly frequented shopping streets in Germany and Austria combined in terms of foot traffic (Engel & Völkers Commercial GmbH, 2013). Both of these facets may go some way

---

13. An anonymous reviewer responded to this statement with "Well, you don't need 'robust comparative evidence' to tell me e.g. that there are more trees in a forest than in a meadow". This simile would be valid if the nature of a 'commercial LL' were *defined* by bearing a lot of signs, which, obviously, it is not (it is defined by being commercial). In other words, a commercial LL is not by definition an accumulation of signs (like a forest is by definition an accumulation of trees and a meadow is a plot of grassland) – it is defined by economic activity, which does not in theory or by any inherent necessity entail more signage than residential activity (for example, even residential buildings may bear a name, just like shops). We are still dealing with an everyday impression or expectation, therefore, that warrants empirical/quantitative validation – albeit perhaps in the line of 'quod erat demonstrandum'.

towards explaining the extra-high LL density (concentrated signaling activity). In the −MULTILINGUALISM district, the surveyed commercial area is the outer perimeter of a large triangular square in which three major roadways intersect ('Am Spitz' – see also Footnote 9). It features an unusually high number of traffic signs, cross lights, and other 'street furniture' (trash cans, phone booths, hydrants, bike racks, etc.) – a fact that already struck us during fieldwork, though no particular reason for this accumulation could be discerned. 'Am Spitz' contains 19 pieces of street furniture per 100m, whereas 'Graben' contains 11, and all other street sections in the ELLViA sample range between 2 and 8 pieces per 100m. Taking all items affixed to street furniture out from the entire sample considerably evens the count, as the rate of items per meter for 'Am Spitz' drops to 5. All other shopping streets also range between 3 and 5 items per m in this new count, which thus seems to be the expected rate of LL items for the typical (Viennese) shopping street. Only 'Graben' remains an outlier, with 9 items per m still. In the new count, all residential streets feature only one item per meter, as the rate at which LL sign readers encounter written language as they move and look along an urban, built-up landscape.

At the time of writing this article, the ELLViA corpus was not yet in its entirety coded for language choice on the individual items, preventing any distributional analyses in this regard. However, based on the field notes, comparisons are possible regarding the frequencies of the *types* of items found, which project implications for a language use analysis down the line.

One instance where some results are already available, and which may thus be adduced to illustrate the considerable potential of the VaLLS approach, concerns the item type of stickers in the ELLViA dataset ($N=4{,}583$ out of 17,109; or about 27% of all items). Schuster (2018) provides a detailed study of those ELLViA stickers that were recorded in the 8th district, which represents the 'young adult' population in the sample (see Section 3.1. above). Her analysis focuses in particular on stickers that were evidently placed illegally ($N=453$ out of 782 stickers total, or 58%). Such stickers feature English language use at a rate of 27% ($N=122$ out of 453) in Schuster's 8th district sub-sample (English in some combination with German: 13%; English only or with other languages: 14%).

An (unpublished) pilot study for the ELLViA project, consisting of over 7,000 Viennese LL items collected in non-ELLViA districts by university students in the course of a seminar, found a similar rate of English language use, of around 30%, on all illegally placed stickers. By contrast, English language use in this fully coded pilot dataset dropped to around 20% on all the other items, outside of those stickers. Taken together, the cited findings predict that, wherever more illegally placed stickers are found in the Viennese LL, we can also expect higher rates of English language use.

As it turns out, the different hypotheses-based ELLViA sub-samples do exhibit different rates of illegally placed stickers, as shown in Table 3.

Table 3. Frequencies of illegally placed stickers in the paired ELLViA districts (labeled according to the variable operationalized)

	LL items per 100m (N)	Illegally placed stickers per 100m (N)	Percentage of illegally placed stickers in all items
−AGE (=young)	291	55	19.0%
+AGE (=old)	292	7	2.4%
+MULTILINGUALISM	252	11	4.5%
−MULTILINGUALISM	456	11	2.5%
+TOURISM	521	23	4.4%
−TOURISM	295	5	1.5%
Total	**352**	**19**	**5.4%**

Table 3 relates the frequency of illegally placed stickers summarily for each district (the ratios of such stickers per all items are quite similar for the paired commercial and residential streets, so these results are not presented separately). We find the highest number as well as percentage of illegally placed stickers per 100m in the district that represents the young adult population (−AGE in Table 2; as mentioned above, this is also the district analyzed in Schuster's study). The percentages recorded in the highly multilingual and the highly tourist districts are both still larger than those in their counterparts (−MULTILINGUALISM and −TOURISM). In the MULTILINGUALISM pairing, however, the *amount* of illegally placed stickers per 100m is equally high ($N=11$), because the −MULTILINGUAL-ISM sample contains many more items per 100m than the +MULTILINGUALISM one ($N=456$ vs. $N=252$). This means that, while LL readers will typically encounter just as many illegally placed stickers in both districts, these stickers make up more of the entire LL in the +MULTILINGUAL setting. Whether and how this affects their perceptual impact is a question that must remain open to discussion at this point. At the same time, it is a question we could not even begin to consider without the precise circumscription of the survey area and the count-all data collection procedure implemented here.

Extrapolating from the findings of Schuster (2018) and the ELLViA pilot discussed above, we can now predict that occurrences of English positively correlate with the frequency of illegally placed stickers in Vienna. In other words, English language use will be heightened in those districts with more illegally placed stickers (−AGE, +TOURISM, and, percentage-wise, +MULTILINGUALISM) compared to their counterparts (+AGE, −TOURISM, −MULTILINGUALISM), in a trend that sup-

ports the relevant hypotheses of where 'more' English would be found in the city (see again Section 3.1.).

While on the surface, illegally placed stickers may appear like a subcultural phenomenon that operates outside of, or even counter-culturally to, the 'sanctioned' LL, a large proportion of them actually exhibit run-of-the-mill promotional interests, constituting 'guerilla marketing'. Thus, Schuster (2018) finds that around two-thirds of the illegally placed stickers in her subsample of the 8th district serve a commercial purpose ('Werbung'), pushing tangible products (such as CDs, books, or fashion) or promoting entertainment events (concerts, movies). This in turn previews a potential positive correlation between English language use and commercial activities, which is also an ELLViA test hypothesis. All in all, then, the sticker analysis exemplarily showcases the kinds of distribution-based, correlative analyses for which a VaLLS approach readily prepares the ground.

5.   Conclusion and outlook

As illustrated in the design and early findings of the ELLViA VaLLS project, the application of variationist principles to LL surveying is a fruitful strategy for generating robust results and original perspectives on the physical distribution of written language in public space. On the basis of well-established tenets (hypothesis-driven stratified sampling, the Principle of Accountability and its concomitant count-all procedure), variationist LL research allows us to explore quantitatively (statistically) the intricate yet systematic ways in which language choice is linked to social factors such as local LL readership configurations and ambient activity types.

Of course, quantitative extraction and analysis of correlations in the data are only ever able to tell part of the story behind sociolinguistic variation and language choice – in the LL and elsewhere. They help us establish a general picture of common behavioral patterns. In order to understand the functionality of these patterns – or, how exactly language use both reflects and shapes social life on the ground – we need to trace the communicative meaning-making activities from which the patterns arise. This may be achieved by integrating quantitative research into a more comprehensive agenda investigating variation as a socio-interactional phenomenon – which is the agenda of the current 'third wave' of variationist sociolinguistics (see Eckert, 2012).

Under the third wave, variation study goes beyond establishing mere correlations between language use and social categories (such as region of origin, socioeconomic status, ethnicity, gender), which has been its traditional, and indeed informative, program. The new, expanded focus is on how variation as such serves

as an agentive, strategic, rhetorical tool that may be used pro-actively to create communicative effects in interaction, such as projections of personas and relationships (Eckert, 2012; Schilling, 2013b).

The third wave is predicated on the very dialogical model of communication I have outlined further above (in Section 3.1.). To recap, this model holds meaning to be jointly achieved, rather than 'transmitted', between interactional participants, through mutual anticipation, negotiation, and interpretation. Hence, the recipiency of language use (including language choice) plays as vital and constitutive a role in communication as its production. A central mechanism of the dialogue between production and recipiency is 'contextualization' (Gumperz, 1982), whereby aspects of the context of language use, including language choice, are indexed (by language producers) and adduced (by recipients) in the comprehensive process of interactional, joint, situated meaning-making.[14]

Adopting a dialogical view in LL research casts any quantitative, distributional patterns of written language choice as physical manifestations of an (albeit mostly asynchronous) interaction between sign authors and sign readers. Authors put forth, and readers infer, messages shaped, among other things, by the featured language choice. In tribute to the complexity of the concomitant meaning-making processes, a comprehensive analysis of these physical manifestations (i.e., LL items) necessitates a mixed-methods design that integrates the systematic (quantitative) extraction of the general scope, shape, and patterning of LL 'dialogue' in a particular setting, with other kinds of research (such as ethnography, discourse analysis, and social psychological language attitude study), in order to establish which social meanings are associated with particular kinds of language use in this very setting, and what kinds of messages may thus arise in the dialogical LL-author-recipient interaction.[15]

Like other methodologies, VaLLS is thus best regarded as a means to an end, and not an end in and of itself, within the overarching program of LL research. It is meant to address known criticism, shortcomings, and desiderata regarding quantitative LL study, so that the latter may better realize its potential as "a prerequisite for LL research which seeks to describe and analyse a given space" under "a

---

14. For general discussion of the dialogical model in relation to third wave variation study, see Soukup (2018); for theoretical modelling and further illustration, see Soukup (2013). Note again and explicitly that 'joint meaning-making' does not necessarily imply nor require a temporally synchronous staging – rather, it refers to 'thinking' or 'projecting' the other, as a function of the inherent dialogicality of *all* communication (whether or not it takes place face-to-face), as notably theorized by Bakhtin (1986 [1952–53]) and explicated by Linell (2017).

15. See Soukup (2015) for a programmatic proposal for interdisciplinary mixed-methods research in third wave variation study, elaborated in the context of spoken language (code-switching between Austrian German dialect and standard in TV discussions).

symbiotic approach, where the quantitative and qualitative approaches feed into one another" (Blackwood, 2015, p. 40). Used systematically and in a rigorous fashion along the lines discussed here, VaLLS arguably gets us that much closer to understanding the sociolinguistic dynamics of LLs – why what's there is there.

Funding

Acknowledgements

References

Amos, H. W. (2016). Chinatown by numbers: Defining an ethnic space by empirical linguistic landscape. *Linguistic Landscape*, 2(2), 127–156. https://doi.org/10.1075/ll.2.2.02amo

Amos, H. W., & Soukup, B. (2020). Quantitative 2.0: Towards Variationist Linguistic Landscape Study (VaLLS) and a standard canon of LL variables. In D. Malinowski & S. Tufi (Eds.), *Reterritorializing Linguistic Landscapes: Questioning boundaries and opening spaces* (pp. 56–66). London: Bloomsbury.

Backhaus, P. (2007). *Linguistic landscapes. A comparative study of urban multilingualism in Tokyo*. Clevedon: Multilingual Matters.

Bakhtin, M. (1986 [1952–53]). The problem of speech genres. In C. Emerson & M. Holquist (Eds.), V. W. McGee (Transl.), *Speech genres and other late essays* (pp. 60–102). Austin: The University of Texas Press.

Ben-Rafael, E., Shohamy, E., Amara, M. H., & Trumper-Hecht, N. (2006). Linguistic landscape as symbolic construction of the public space. The case of Israel. *International Journal of Multilingualism*, *3*(1), 7–30. https://doi.org/10.1080/14790710608668383

Ben-Rafael, E., Shohamy, E., & Barni, M. (2010). Introduction: An approach to an 'ordered disorder'. In E. Shohamy, E. Ben-Rafael, & M. Barni (Eds.), *Linguistic landscape in the city* (pp. xi–xxviii). Clevedon: Multilingual Matters. https://doi.org/10.21832/9781847692993-002

Berns, M., de Bot, K., & Hasebrink, U. (Eds.). (2007). *In the presence of English. Media and European youth*. New York: Springer. https://doi.org/10.1007/978-0-387-36894-8

Blackwood, R. J. (2011). The linguistic landscape of Brittany and Corsica. A comparative study of the presence of France's regional languages in the public space. *Journal of French Language Studies*, *21*(2), 111–130. https://doi.org/10.1017/S0959269510000281

Blackwood, R. (2015). LL explorations and methodological challenges. Analysing France's regional languages. *Linguistic Landscape*, *1*(1/2), 38–53. https://doi.org/10.1075/ll.1.1-2.03bla

Blackwood, R. J., & Tufi, S. (2015). *The linguistic landscape of the Mediterranean. French and Italian coastal cities*. Basingstoke: Palgrave Macmillan. https://doi.org/10.1057/9781137314567

Brizic, K. (2013). Grenzenlose Biografien und ihr begrenzter (Bildungs-)Erfolg. Das Thema der sozialen Ungleichheit aus der Perspektive eines laufenden soziolinguistischen Forschungsprojekts. In A. Deppermann (Ed.), *Das Deutsch der Migranten* (pp. 223–242). Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110307894.223

Brizic, K., & Hufnagl, C. L. (2011). "Multilingual Cities" Wien. Bericht zur Sprachenerhebung in den 3. und 4. Volksschulklassen. Vorläufiger Bericht für die teilnehmenden Schulen. Retrieved on June 21, 2019, from https://ec.europa.eu/migrant-integration/?action=media.download&uuid=FBD38E84-0854-9B35-C7367CF1FD90508D

Bruyèl-Olmeido, A., & Juan-Garau, M. (2009). English as a lingua franca in the linguistic landscape of the multilingual resort of S'Arenal in Mallorca. *International Journal of Multilingualism*, *6*(4), 386–411. https://doi.org/10.1080/14790710903125010

Cenoz, J., & Gorter, D. (2006). Linguistic landscape and minority languages. *International Journal of Multilingualism*, *3*(1), 67–80. https://doi.org/10.1080/14790710608668386

Chambers, J. K. (2008). *Sociolinguistic theory*. 2nd edition. Oxford: Wiley-Blackwell.

Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford: Oxford University Press.

Eckert, P. (2012). Three waves of variation study. The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology*, *41*, 87–100. https://doi.org/10.1146/annurev-anthro-092611-145828

Edelman, L. (2014). The presence of minority languages in linguistic landscapes in Amsterdam and Friesland (the Netherlands). *International Journal of the Sociology of Language*, *228*, 7–28.

Engel & Völkers Commercial GmbH (2013). Frequenzzählung 2013 / Frequency survey 2013. *Retail Services*. Retrieved on March 19, 2015, from https://www.engelvoelkers.com/pdfs/research/Frequenzzaehlung_2013_Top73.pdf

Erickson, F. (1986). Listening and speaking. In D. Tannen & J. E. Alatis (Eds.), *Languages and linguistics. The interdependence of theory, data, and application* (pp. 294–319). Washington, DC: Georgetown University Press.

Fasold, R. (1990). *The sociolinguistics of language*. Oxford: Blackwell.

Goffman, E. (1959). *The presentation of self in everyday life*. New York: Doubleday.

Gorter, D. (Ed.). (2006). *Linguistic landscape. A new approach to multilingualism*. Clevedon: Multilingual Matters. https://doi.org/10.21832/9781853599170

Gorter, D. (2013). Linguistic landscapes in a multilingual world. *Annual Review of Applied Linguistics*, *33*, 190–212. https://doi.org/10.1017/S0267190513000020

Grau, M. (2009). Worlds apart? English in German youth cultures and in educational settings. *World Englishes*, *28*(2), 160–174. https://doi.org/10.1111/j.1467-971X.2009.01581.x

Gumperz, J. J. (1982). *Discourse strategies*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511611834

Gumperz, J. J. (2001). Interactional sociolinguistics. A personal perspective. In D. Schiffrin, D. Tannen, & H. E. Hamilton (Eds.), *The handbook of discourse analysis* (pp. 215–228). Malden: Blackwell.

Guy, G. R. (1993). The quantitative analysis of linguistic variation. In D. R. Preston (Ed.), *American dialect research* (pp. 223–249). Philadelphia: John Benjamins. https://doi.org/10.1075/z.68.11guy

Hilgendorf, S. K. (2007). English in Germany. Contact, spread and attitudes. *World Englishes*, *26*(2), 131–148. https://doi.org/10.1111/j.1467-971X.2007.00498.x

Hofmann, D. (2002). Do you understand Denglish? Eine Umfrage zum Anglizismenverständnis. In R. Hoberg (Ed.), *Deutsch – Englisch – Europäisch. Impulse für eine neue Sprachpolitik* (pp. 236–245). Mannheim: Dudenverlag.

Huebner, T. (2006). Bangkok's linguistic landscapes: Environmental print, codemixing and language change. *International Journal of Multilingualism*, *3*(1), 31–51. https://doi.org/10.1080/14790710608668384

Kachru, B. B. (1992). Teaching World Englishes. In B. B. Kachru (Ed.), *The other tongue. English across cultures*, 2nd edition (pp. 355–366). Urbana: University of Illinois Press.

Kelly-Holmes, H. (2000). Bier, parfum, kaas: Language fetish in European advertising. *European Journal of Cultural Studies 3*(1), 67–82. https://doi.org/10.1177/a010863

Kelly-Holmes, H. (2005). *Advertising as multilingual communication*. New York: Palgrave Macmillan. https://doi.org/10.1057/9780230503014

Kiesling, S. (2011). *Language variation and change*. Edinburgh: Edinburgh University Press.

Labov, W. (1966). *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.

Labov, W. (1969). Contraction, deletion, and inherent variability of the English copula. *Language*, *45*(3), 715–762. https://doi.org/10.2307/412333

Leppänen, S., Pitkänen-Huhta, A., Nikula, T., Kytölä, S., Törmäkangas, T., Nissinen, K., Kääntä, L., Räisänen, T., Laitinen, M., Koskela, H., Lähdesmäki, S., & Jousmäki, H. (2011). *National survey on the English language in Finland. Uses, meanings and attitudes*. Helsinki: VARIENG. Available from http://www.helsinki.fi/varieng/series/volumes/05/

Levy, P. S., & Lemeshow, S. (2008). *Sampling of populations. Methods and applications*, 4th edition. Hoboken: Wiley. https://doi.org/10.1002/9780470374597

Linell, P. (2017). Dialogue, dialogicality and interactivity. *Language and Dialogue*, *7*(3), 301–335. https://doi.org/10.1075/ld.7.3.01lin

Magistrat Wien (2014). *Statistisches Jahrbuch der Stadt Wien – 2014*. Retrieved on February 19, 2015, from https://www.wien.gv.at/statistik/publikationen/uebersicht-pub.html

Milroy, L. & Gordon, M. (2003). *Sociolinguistics. Method and interpretation*. Malden: Blackwell.

Piller, I. (2003). Advertising as a site of language contact. *Annual Review of Applied Linguistics*, *23*, 170–183. https://doi.org/10.1017/S0267190503000254

Piritidis, K. (2014). The linguistic landscape of Vienna's 'Westbahnstraße'. *A comparison of methods*. Unpublished Mag. phil. thesis, University of Vienna.

Pitkänen-Huhta, A., & Nikula, T. (2013). Teenagers making sense of their foreign language practices. Individual accounts indexing social discourses. In P. Benson & L. Cooker (Eds.), *The applied linguistic individual. Sociocultural approaches to autonomy, agency and identity* (pp. 101–120). Sheffield: Equinox.

Prensky, M. (2001). Digital natives, digital immigrants. Part 1. *On the Horizon*, *9*(5), 1–6. https://doi.org/10.1108/10748120110424816

Reichl, S. (2009). *Cognitive principles, critical practice. Reading literature at university*. Göttingen: V&R unipress.

Rode, P., Schier, H., Giffinger, R., & Reinprecht, C. (2010). *Soziale Veränderungsprozesse im Stadtraum. Wiener Sozialraumanalyse mit Vertiefung in acht ausgewählten Stadtvierteln*. Vienna: Stadtentwicklung Wien, Magistratsabteilung 18.

Sankoff, D. (2005). Problems of representativeness/Probleme der Repräsentativität. In U. Ammon, N. Dittmar, K. J. Mattheier, & P. Trudgill (Eds.), *Sociolinguistics/Soziolinguistik*. Vol. 2, 2nd edition (pp. 998–1002). Berlin: Walter de Gruyter.

Schiffrin, D. (1994). *Approaches to discourse*. Malden: Blackwell.

Schilling, N. (2013a). *Sociolinguistic fieldwork*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511980541

Schilling, N. (2013b). Investigating stylistic variation. In J. K. Chambers & N. Schilling (Eds.), *The handbook of language variation and change*, 2nd edition (pp. 325–349). Oxford: Wiley. https://doi.org/10.1002/9781118335598.ch15

Schuster, C. (2018). Analyzing stickers in the linguistic landscape of Vienna, Austria. Unpublished MA thesis, University of Vienna.

Seidlhofer, B. (2011). *Understanding English as a Lingua Franca*. Oxford: Oxford University Press.

Siricharoen, A. (2016). Multilingualism in the linguistics landscape of the Faculty of Arts, Chulalongkorn University, Thailand. *Manusya: Journal of Humanities, Special Issue 22*, 12–25. https://doi.org/10.1163/26659077-01903002

Soukup, B. (2009). *Dialect use as interaction strategy. A sociolinguistic study of contextualization, speech perception, and language attitudes in Austria*. Vienna: Braumüller.

Soukup, B. (2013). Austrian dialect as a metonymic device: A cognitive sociolinguistic investigation of Speaker Design and its perceptual implications. *Journal of Pragmatics*, *52*, 72–82. https://doi.org/10.1016/j.pragma.2012.12.018

Soukup, B. (2015). Mixing methods in the study of language attitudes: Theory and application. In A. Prikhodkine & D. R. Preston (Eds.), *Responses to language varieties. Variability, processes and outcomes* (pp. 55–84). Amsterdam: John Benjamins. https://doi.org/10.1075/impact.39.03sou

Soukup, B. (2016). English in the linguistic landscape of Vienna, Austria (ELLViA): Outline, rationale, and methodology of a large-scale empirical project on language choice on public signs from the perspective of sign-readers. *Views*, *25*, 1–24.

Soukup, B. (2018). Contextualizing the third wave in variationist sociolinguistics: On Penelope Eckert's (2018) *Meaning and Linguistic Variation*. *Views*, *27*, 51–66.

Spolsky, B., & Cooper, R.L. (1991). *The languages of Jerusalem*. Oxford: Clarendon Press.

Statistik Austria (2013). *Census 2011 Wien. Ergebnisse zur Bevölkerung aus der Registerzählung.* Vienna: Statistik Austria.

Statistik Austria (2019). Durchschnittsalter der Pensionsneuzuerkennungen in der gesetzlichen Pensionsversicherung 1970 bis 2018. Retrieved on October 20, 2019, from http://www.statistik.at/wcm/idc/idcplg?IdcService=GET_PDF_FILE&RevisionSelectionMethod=LatestReleased&dDocName=058368

Tagliamonte, S.A. (2006). *Analysing sociolinguistic variation*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511801624

Tagliamonte, S.A. (2012). *Variationist sociolinguistics. Change, observation, interpretation.* Malden: Wiley-Blackwell.

Tannen, D. (1989). *Talking voices*. Cambridge: Cambridge University Press.

Tannen, D. (2004). Interactional sociolinguistics/Interaktionale Soziolinguistik. In U. Ammon, N. Dittmar, K.J. Mattheier, & P. Trudgill (Eds), *Sociolinguistics/Soziolinguistik*, Vol.*1*, 2nd edition (pp. 76–88). Berlin: Walter de Gruyter.

Thompson, S.K. (2012). *Sampling*. 3rd edition. Hoboken: Wiley. https://doi.org/10.1002/9781118162934

Walker, J. (2010). *Variation in linguistic systems*. New York: Routledge.

Weinreich, U., Labov, W., & Herzog, M.I. (1968). Empirical foundations for a theory of language change. In W.P. Lehmann & Y. Malkiel (Eds.), *Directions for historical linguistics* (pp. 95–188). Austin: University of Texas Press.

Wirtschaftsagentur Wien (2017). Richtlinie des Förderprogramms Nahversorgung 2017. Retrieved on June 21, 2019, from https://cockpit.wirtschaftsagentur.at/Kontext/Webservice/SecureFileAccess.aspx?fileguid=%7Bf425dc6f-8077-45f7-a9e7-884494d7c54a%7D

Abstrakt

Dieser Beitrag befasst sich mit der bis dato ungelösten Problematik der systematischen Absteckung eines Untersuchungsgebiets für umfangreiche, quantitative Studien von Sprachlandschaften (linguistic landscapes/LLs). Eine Strategie des ‚hypothesengeleiteten, stratifizierten Samplings' wird vorgestellt, unter der ein Untersuchungsgebiet, in mehreren verschränkten Schritten, auf Basis der Konfiguration eines lokalen LL-Publikums ausgewählt wird, und zwar im Hinblick auf dessen Alter, Multilingualismus und touristische Aktivität, sowie auf Umgebungstyp (kommerziell oder residenziell). Die Logik dieser Strategie ist von der variationistischen Soziolinguistik abgeleitet, sodass das Unterfangen als ‚variationistische Sprachlandschaftsforschung' ('Variationist Linguistic Landscape Study'/ ‚VaLLS') bezeichnet werden kann. Die Details eines entsprechenden Studiendesigns und seine konkrete Umsetzung, sowie die Implikationen für die LL-Forschung im Allgemeinen, werden anhand des Projekts ‚ELLViA – Englisch in der Wiener Sprachlandschaft' illustriert und diskutiert. Letztlich wird gezeigt, dass die Anwendung moderner variationistischer Prinzipien und Methodologien die wissenschaftliche Stringenz und Qualität der quantitativen LL-Forschung, die diesbezüglich laufend in der Kritik stand, erheblich verbessern kann.

Address for correspondence

Barbara Soukup
Department of German Studies
University of Vienna
Universitätsring 1
A-1010 Vienna
Austria

barbara.soukup@univie.ac.at