

The measurement of ‘language attitudes’ – a reappraisal from a constructionist perspective

Barbara Soukup

*University of Vienna*¹

As handbooks on the social psychological study of attitudes routinely profess, over the past hundred years, ever since the inception of the field, the most commonly used method in attitude measurement has been the use of rating scales (e.g. Schwarz 2008).² Such scales can take a variety of forms, though the three classic types are those developed in the works of Thurstone (e.g. 1928), Likert (e.g. 1932), and Osgood and associates (e.g. Osgood, Suci and Tannenbaum 1957) – (for further discussion see Himmelfarb 1993; Krosnick, Judd and Wittenbrink 2005). Indeed, Osgood et al.’s ‘semantic differential’ attitude measure ‘is the foundational technique used most often in research today’ (Krosnick, Judd and Wittenbrink 2005: 33). Under this technique, a series of selected antonymic adjectives are placed at opposite ends of scales (traditionally of seven increments). Informants then evaluate a given stimulus by placing a checkmark on each scale, with the instruction that the closer they tick to either end, the more they indicate the respective adjective pole to apply to the stimulus. Informants’ attitudes are then computed via a compilation of scores from the scales.

The popularity of the semantic differential technique is largely explicable by its simplicity in terms of design and administration. Unlike Thurstone and Likert scales, which require extensive prior calibration and must be designed anew for each research context, Osgood attitude measurements typically involve adjectives ‘that are very general and heavily saturated with evaluative meaning’ (Himmelfarb 1993: 57), as well as thoroughly researched from all possible an-

¹ I cordially thank the editors of this volume, Stef Grondelaers and Tore Kristiansen, for their invaluable feedback on earlier versions of this paper, as well as the participants and audience of the panel on ‘New Approaches in the Study of Language Attitudes’ at the 18th Sociolinguistics Symposium in Southampton (2010) for their comments on its first incarnation.

² although it seems that recently, research using response-latency measures such as the Implicit Association Test (Greenwald, McGhee and Schwartz 1998) has also been gaining some notable momentum.

gles over the past decades, so that presumably little preparation and prior scaling are needed.³ This is why the scales have been labeled ‘the attitude researcher’s “ever-ready batteries”’ (Himmelfarb 1993: 57). In addition, the facts that scales in general easily lend themselves to presentation in questionnaire format, and are quite straightforward in explanation, facilitate the economical polling of large informant samples.

Given these advantages, it is not surprising that the semantic differential scale has also been the dominant response scheme in research on ‘*language attitudes*’, particularly within the ‘speaker assessment’ paradigm (see also Garrett 2005, 2010). Indeed, combining some form of the ‘matched-guise technique’ (Lambert, Hodgson, Gardner and Fillenbaum 1960), whereby informants listen to and assess different speech samples, with a questionnaire on the basis of Osgood scales seems to have become the standard in the (quantitative) measurement of attitudes towards variation in language use. Thus, out of fifty-three matched or verbal guise studies that appeared over the years 2000–2010, thirty-nine (74%) applied the semantic differential (while seven used Likert and four some other form of scales, and only three employed a different response format altogether).⁴

However, over the years, attitude research along these lines has also become the target of much criticism, particularly from the vantage point of social constructionism (see e.g. Gergen 2008; Potter and Wetherell 1987). The main thrust has been that the experimental method involved generates only a poor image of people’s contextually situated, differentiated, and variable evaluative practices, but also, more fundamentally, that the very search for stable, measurable, incorporated ‘attitudes’ is essentially unwarranted.

The purpose of my present paper, then, is to reflect upon such criticism and its implications for present-day ‘language attitude’ research, particularly as regards its empirical methodology. I begin by addressing the more fundamental

³ though adaptation to research context is still advised – see e.g. Garrett (2010) for discussion; see also Osgood, Suci and Tannenbaum (1957: e.g. ch.3).

⁴ Studies were compiled via the LLBA database (Linguistics and Language Behavior Abstracts – ProQuest, August 2011), using the search terms ‘matched guise’ and ‘verbal guise’ in citations and abstracts. A total of eighty-eight relevant articles were found, but only fifty-three of these were retrievable for closer scrutiny. Note that these numbers are sure to under-shoot the total of ‘speaker evaluation’ studies published over the past decade.

There actually appears at times some imprecision in the labeling of the measuring scales in some studies, in the sense that uni-polar adjective scales tend to be labeled as ‘Likert’, although the fundamental principle of the latter is the response to standardized attitudinal statements – see e.g. Himmelfarb (1993); Krosnick, Judd and Wittenbrink (2005) for reference. I am here including uni-polar adjective scales in the semantic differential count, which is arguably closer in principle.

issue cited above, regarding the nature of 'attitudes' and the contingencies of attitude measurement via scales, first, within social psychology at large (the 'mother discipline'), and then within 'language attitude' research in particular (as its 'daughter discipline'). I then address further points of constructionist criticism of quantitative 'language attitude' study, such as the issues of the treatment of context and of response variability. The reassuring upshot of my discussion will be that the quantitative methodology commonly used in 'speaker assessment' research (and particularly the use of semantic differential scales) is defensible even within a constructionist paradigm, although for this we need to reposition our concepts and approach in some fundamental ways.

WHAT WE ARE MEASURING WHEN WE ARE MEASURING '(LANGUAGE) ATTITUDES'

Like most attitudinal measurements, the semantic differential is typically applied in social psychological research at large in the form of multi-item scaling, or, using 'a cluster of several differently worded items that focus on the same target. The item scores for the similar questions are summed, resulting in a total score [...] Thus, multi-item scales maximize the stable component that the items share and reduce the extraneous influences unique to the individual items' (Dörnyei 2007: 103–104). As Krosnick, Judd and Wittenbrink (2005: 33) put it succinctly, the idea behind all Thurstone, Likert, and Osgood scaling is 'the administration of a large set of questions to measure a single attitude'.

The underlying assumption here, of course, is that of the existence of a measurement 'target' – of some coherent entity of an inner state: an 'attitude'. This assumption seems to be upheld in much of today's social psychological academic discourse, although current definitions do somewhat relativize claims about the stability and durability of attitudes based on research findings regarding attitude variation and change. Thus, Eagly and Chaiken's seminal definition of 'attitude' as '*a psychological tendency that is expressed by evaluating a particular entity with some degree of favor or disfavor*' (Eagly and Chaiken 2005: 745, italics in original) chooses the term 'tendency' over the options of 'state' and 'disposition' as a middle ground between absolute temporariness and absolute durability, but retains the implication of a 'latent property' or 'inner state' that people acquire and that gives rise to expressions of evaluative judgments (see also Eagly and Chaiken 1993).

It is, however, this very notion that evaluative responses are supposedly linked with and based on stable, underlying states of mind which social constructionist scholars have come to criticize as problematic at best (e.g. Potter and Wetherell 1987), and as ‘wholly gratuitous’ at worst (Gergen 2008: 335). As Potter and Wetherell (1987) pointed out in their often-cited proposal for a discourse analytic approach to social psychology, traditional conceptualizations of attitudes as enduring, measurable mental entities run into explanatory difficulties when faced with the facts of variability and changeability in informants’ attitudinal responses across or even within situational contexts, but also in terms of inconsistencies between measured attitudes and behavioral outcomes. In fact, the accumulated empirical evidence on these counts has given rise to alternative theoretical and methodological tacks even within the ‘stable-entity approach’. One has been to devise multidimensional measurement scales to capture more of the complexity of underlying attitudes, and then to have this complexity account for variability (see Potter and Wetherell 1987 for discussion, who cite McConahay 1985 as an example). Arguably, however, ‘[t]he variability in people’s discourse cannot be explained merely as a product of a more complex multifaceted attitudinal structure which a more complex scale can assess, because the views expressed vary so radically from occasion to occasion’ (Potter and Wetherell 1987: 53). Some approaches furthermore propose a host of intervening variables between attitudes and measurement outcome or behavior (see again Potter and Wetherell 1987, with reference to Fishbein and Azjen 1975; see also Bassili and Brown 2005; Bohner and Dickel 2011 for review). However, as Potter and Wetherell (1987: 54) contend, such proposals run ‘the danger of massive post hoc interpretation’ just so as to be able to uphold the concept of attitudes as underlying entities. Thus, ‘[g]iven enough modifying variables huge flexibility in response can be explained.’ But in the authors’ view, ‘there must come a point when it is no longer useful to continue stressing the underlying attitude’ (Potter and Wetherell 1987: 54).

Under the pressure of empirical evidence regarding the context-dependence and variability of attitudinal responses, attitudes have thus more recently been theorized ‘not as enduring personal dispositions [...] but rather as evaluative judgments that are constructed in the situation based on currently accessible information’ (Bohner and Dickel 2011: 393, with reference to proposals by Schwarz 2007). Such a constructionist perspective is also evident in Potter and Wetherell’s (1987) proposal that attitudinal scores elicited via scales are in fact ‘a specific linguistic formulation tuned to the context at hand’ (Potter and Weth-

erell 1987: 45), and in Potter's (1998) conceptualization of attitude expression in terms of 'evaluative practice' or 'activity'.

However, extreme constructionist positions that would seem to abandon any assumption of stable attitudes in favor of contextually relative, ad hoc evaluative construals have in turn been heavily criticized. Thus, Bassili and Brown (2005: 566) suggest that '[t]o conclude that an attitude that changes with shifting contexts is not based on a stable representation is akin to concluding that a flag that changes direction with shifting winds is not attached to a flagpole'. Similarly, Eagly and Chaiken (2005: 746) maintain that '[t]he main reason why some investigators have concluded that most, if not all, attitudes are unstable, constantly emerging anew in specific situations, is that they have equated variability in the expression of attitudes with variability in the evaluative tendency that constitutes attitude', or, in other words, '[t]his attitudes-as-constructions position [...] conflates variability in attitudinal responses with variability in attitude itself'.

All in all, the ongoing debate in social psychology about the existence, definition, granularity, and role of stored mental representations of evaluation (attitudes) is reminiscent of the age-old mentalist vs. behaviorist discussion related for example in Agheysi and Fishman (1970), which pivots on the question of whether or not 'attitude' is equivalent with 'attitudinal response'.⁵ It is certainly impossible to resolve this issue here. However, now that we have seen how the 'mother discipline' struggles with the conceptualization of 'attitude', the question suggests itself of how *language attitude* research commonly defines the relation between attitudinal measurement, its target, and its outcome, in the context of its typical quantitative, scale-based elicitations of 'speaker assessments'.

One central aspect to note here, then, is that virtually from the beginning, quantitative speaker assessment research did not actually follow along with social psychological attitude study and its fundamental methodological principle of using multi-item scales to derive a single attitudinal score. Thus, the basic analysis in Lambert et al.'s (1960) debut of the matched-guise technique does not initially provide any such single score, but rather presents the results for each of their fourteen personality trait scales in turn (though an overall score is used later for correlation with other psychological measures). Lambert (1967) subsequently reports on a study by Preston (1963) in which *three* dimensions of personality judgment were applied to the set of scales used ('competence', 'per-

⁵ Relatedly, see also e.g. Farbrigar, MacDonald and Wegener (2005) for review and discussion of the widely debated question of whether or not attitudes have subcomponents (such as the well-known triad of cognitive, affective, and conative components).

sonal integrity', and 'social attractiveness'). This originally rather ad hoc grouping of items has since been replicated in much 'language attitude' research. Zahn and Hopper (1985) derived evidence for the multidimensional character of speaker assessments more methodically from a factor analysis of thirty semantic differential scales, which yielded their own three rating categories ('superiority', 'attractiveness', 'dynamism').

Overall, it seems to be received knowledge in the 'language attitude' literature that it does not make much sense to boil the results of speaker assessment experiments down to one single attitude score – because what would this score even tell us? The bulk of (socio)linguistic research has shown that people's dealings with language are much more subtle than sweeping statements of 'favor' or 'disfavor' may ever hope to capture. Rather, there are likely to be several lines of consideration along which listeners typically assess speakers and their language use, which needs to be reflected in the selection and analysis of scale items of respective experiments. However, something that is never mentioned in this context is that, by virtue of this very fact, most 'language attitude' studies are actually not really measuring language *attitudes*.

Recall that in social psychology at large, 'attitude' connotes very particularly a 'degree of favor or disfavor' (see the definition by Eagly and Chaiken 2005 above) – an *evaluation*, the essence of which is the distinction 'good–bad' or 'like–not like' (Eagly and Chaiken 2005; Heise 1970; Krosnick et al. 2005). Consequently, social psychological attitude scales are commonly built to capture this very evaluative aspect – *and this aspect only*. Now, 'language attitude' scales do usually include some measures of 'pure' evaluation (do speakers sound rather 'nice' or 'awful'? – see Zahn and Hopper 1985), but are typically designed so as to check, in addition, things like whether informants hold speakers to sound 'educated' or 'uneducated', 'rich' or 'poor', 'active' or 'passive', 'strong' or 'weak' (see again Zahn and Hopper 1985), or even 'tall', 'entertaining', or 'ambitious' (Lambert et al. 1960), depending on their language use. Clearly, some of these adjective items do not have any objective, decontextualizable positive or negative (i.e. 'evaluative') character (see also Garrett 2010; Lambert et al. 1960). Indeed, ultimately, even sounding 'simple-minded' can be a good thing in certain situations (such as when trying to charm customers into buying products – see Soukup 2011, in the context of Southern American English). It is in this respect, then, that some of the scales used in 'language attitude' research capture something other than 'attitudes' in the strict social psychologi-

cal sense. (Hence my use of inverted commas and my preference for the term 'speaker assessment' over the more common 'speaker evaluation'.)

The point I am making here may seem arcane at first, but it does have important implications. For one, it may go some way towards explaining why 'language attitude' research has been widely ignored by its mother discipline, social psychology at large (a fact lamented by Preston in 2009, who mentioned that the 2005 landmark *Handbook of Attitudes* by Albarracín, Johnson, and Zanna does not include a single reference to work on 'language attitudes'). But what's more, I argue that it actually and quite logically preempts the kind of social constructionist criticism of scale-based attitude measurement that pivots on the accusation of unwarranted attachment to monolithic evaluative dispositions. Such criticism has already been taken to imply, by extension, that 'experimentation with the matched-guise technique [...] should be giving way to discourse-analytic studies of language attitudes' (Hyrkstedt and Kalaja 1998: 348, with reference to Kalaja 1997). If, however, quantitative 'language attitude' studies do not even pretend to be eliciting one comprehensive, underlying attitudinal score, this, for one, cannot be grounds on which to dismiss them.

I have of course still left open the question of what it is, then, that traditional, scale-based 'speaker assessment' research really elicits. To work this out, it pays to return to Osgood et al.'s original presentation of the semantic differential from 1957. Over time, their scaling method has become synonymous with attitude elicitation; however, the title of their book is in fact *The Measurement of Meaning*, not: ... 'of Attitudes'. Indeed, the purpose of their endeavor was to subject meaning to quantitative analysis; defining meaning as 'that process or state in the behavior of a sign-using organism which is assumed to be a necessary consequence of the reception of sign-stimuli and a necessary antecedent for the production of sign-responses' – a cognitive 'representational mediation process' (Osgood, Suci and Tannenbaum 1957: 9). 'Semantic differentiation', then, implies 'the successive allocation of a concept to a series of descriptive scales defined by polar adjectives, these scales selected so as to be representative of the major dimensions along which meaningful processes vary' (p. 31). 'Difference in the meaning between two concepts is then merely a function of the differences in their respective allocations within the same space' (p. 26). Or, put more simply, the meaning of a concept can be expressed in terms of its loadings on the bipolar adjective items bracketing the semantic differential scales.

Based on extensive research, Osgood et al. furthermore extracted three fundamental 'factors' or 'dimensions' to which much variance in meaning assess-

ments can be reduced, which they call ‘evaluation’, ‘potency’, and ‘activation’ (typified by the items ‘good–bad’, ‘powerful–powerless’, and ‘fast–slow’ respectively; Heise 1970). The reason for which Osgood et al.’s scaling method has become closely associated with attitudinal research is their *post hoc* proposal that the ‘evaluation’ dimension of meaning measurement is essentially equivalent with the notion of ‘attitude’. At the same time, however, Osgood et al. did not suggest that the three dimensions they identified are exhaustive or universal, nor that the evaluative dimension must take any sort of precedence. On this basis, it can now be argued that even if we follow social constructionists in rejecting the existence of underlying, single attitudes and the relevance of the quest, this does not ipso facto discredit the semantic differential as an empirical method. Rather, we can choose to refocus on its original purpose, and harness it for the exploration of *meaning* beyond the purely (merely) evaluative. ‘Language attitude’ research, then, actually seems to have a long tradition of applying the semantic differential according to its original intent – locating speakers and their language production within the ‘semantic space’ made up by what we would most likely call the *social meanings* associated with linguistic communication in a given context. For any given study, it may turn out that this social meaning space is indeed largely configured along the lines of the three dimensions of ‘superiority’, ‘attractiveness’, and ‘dynamism’ (Zahn and Hopper 1985), or that there are many more directions of pull, or that dimension reduction is not particularly informative at all (as seems to be one upshot of Potter and Wetherell 1987).

Again, as mentioned further above, one important benefit of calling the overarching endeavor now by its ‘real’ name, *the measurement of the social meaning of linguistic variation*, instead of insisting on measuring ‘language attitudes,’ is that this makes clear why some of the criticism leveled against the field from a social constructionist perspective is actually not warranted. At the same time, it pushes us ‘language attitude’ scholars to finally let our words follow our deeds and to stop obstinately trying to tie our interpretations of findings from scale-based speaker assessments back to the social psychological notion of ‘attitude,’ with its restriction to underlying purely evaluational entities, when what we are finding really goes beyond. In my opinion, this kind of self-imposed disciplinary submissiveness has long outlived its scaffolding purpose, and has been at the root of much terminological and conceptual fuzziness, as well as of the lack of rigorous theorizing, that seem to have hampered scholarly reviews and studies of ‘language attitudes’ to this day.

'SOCIAL MEANINGS OF VARIATION' IN APPLICATION

There is another important affordance that derives from the above-proposed conceptualization of 'language attitudes,' as elicited via scale-based speaker assessment, in terms of 'social meanings of variation.' This is that it may actually counter the opinion commonly held under a social constructionist perspective that quantitative 'language attitude' research is hard to justify these days for the reason that findings are difficult to apply to real-life situations (Hyrkstedt and Kalaja 1998: 346).

After all, the fact that the social meaning of variation plays a central role in sociolinguistic processes anywhere from diachronic language change to interactional persona-management is nowadays fundamentally undisputed, and the topic of a rapidly expanding body of research (see e.g. Kristiansen 2009; Coupland 2007 respectively). Take for example current studies of the phenomenon of 'Speaker Design' (Schilling-Estes 2002), or the way in which language users may harness the linguistic variants and varieties in their repertoire proactively to create interactional identities and alignments. They do this by navigating what Eckert (2008) has called the 'indexical field' of linguistic variants – networks of social ideologies and social meanings associated with certain kinds of language use. The choice of a certain variant in an interaction indexes such associations as relevant for inferencing utterance meaning, giving rise to corresponding interpretations regarding identity projections and participant alignments (see Gumperz' 1982 notion of 'contextualization'). Thus, Austrian participants in a TV discussion show have been found to shift from Austrian standard ('Hochsprache') into Bavarian-Austrian dialect when portraying a political opponent as ignorant and coarse, drawing on (i.e. 'contextualizing their utterances with') corresponding social meanings that are commonly associated with dialect use in Austria (see Soukup 2009).

The fact that Austrian listeners are actually likely to realize these associations and come up with matching interpretations of the Speaker Design (rhetorical shifts into dialect) can be tested via a speaker assessment experiment in which Austrian informants comparatively judge speakers' uses of Austrian standard and dialect. This is because, arguably, what informants are asked to do in the experiment is similar to what an audience is called upon to do when watching the TV show – to make sense of two linguistic varieties in juxtaposition, by drawing on the social meanings associated with these. If the social meanings elicited in the experiment are found to be similar to the ones needed for discus-

sion participants' rhetorical moves of negative portrayal (i.e., that speaking in dialect makes a speaker sound less intelligent and sophisticated than speaking in the standard), this is then evidence for the fact that the Speaker Design will be communicatively successful on TV (and see Soukup 2009 for a data series that yields exactly this result).

Put more generally again, speaker assessment experiments in which listeners comparatively judge speakers' uses of different styles can arguably support interactional findings regarding Speaker Design (the rhetorical use of style-shifting), by yielding the relevant activation patterns in the indexical fields of social meanings surrounding the types of speech under investigation. Ultimately, such potential for application to the investigation of actual, real-life behavior should provide sufficient evidence for the usefulness of speaker assessment elicitation based on semantic differential scales and the 'measurement of (social) meaning'.

FURTHER IMPLICATIONS: RECONCEPTUALIZING SPEAKER ASSESSMENT EXPERIMENTS

Note, then, that current research on Speaker Design and the agentic use of styles for interaction management is very much anchored in constructionist conceptualizations of social life as a function of emergent, contextually situated/relative meaning-making activity (see also Schilling-Estes 2002). Proposals for the application of speaker assessment experiments to investigations of Speaker Design therefore entail that the experimental methodology per se be also epistemologically compatible with a constructionist perspective. This still does not seem to be the case if we, while replacing the notion of 'attitude' with that of 'social meaning', nevertheless conceive of experiments themselves in terms of objective, sterile fact-finding missions, as under a positivist tradition (see the criticism in Giles and Coupland 1991; Potter and Wetherell 1987). Rather, we need to adopt a more modern perspective under which responses on speaker assessment tasks, although artificially induced, are the record of emergent, contextually situated meaning-making activity of the same nature as other types of human social interaction (e.g. everyday conversation). In other words, an experiment should be regarded as a 'discursive event' (Giles and Coupland 1991: 58) in and of itself, in which 'evaluative practice' (Potter 1998) is taking place – albeit under certain characteristic conditions.

On this view, it is the inherent contextual situatedness of *all* communicative events (see e.g. Gumperz 1982) that explains why situational parameters are so vital for the meaning outcome in speaker assessment experiments, just as in other types of interaction. Indeed, time and again it has been found that contextual factors have an important bearing on the recorded results (see e.g. Cargile et al. 1994 for discussion; see also Bohner and Dickel 2011 for discussion within social psychology at large). But instead of regarding and lamenting this as a confounding fact, and trying, but inevitably failing, to keep contextual factors 'out of' an experiment so as to construct some highly general, abstract, underspecified results, it seems more productive to take a proactive approach and match the experiments' situational parameters with some ulterior, concrete context of interactional activity so that the findings can then be applied there. Thus, for example, in the data series briefly mentioned above (from Soukup 2009), the speaker assessment task whose findings were applied to the investigation of Speaker Design in a TV discussion was actually designed so that the contextual parameters obtaining in experiment and interactional data matched quite well: the experimental protocol was aligned to the TV show setting in terms of the introduction and framing of the task (rating 'public speakers' anonymously), the text used (an argumentative statement), and some of the questions asked of the informants (see Soukup 2009, 2010 for details). The meaning-making activities in the experiment (carried out by the informants) and the TV discussion show (carried out by the audience at home) regarding the assessment of the use of different styles (Austrian dialect and standard) could therefore be assumed to have taken place in similar socio-situational contexts, validating application of findings from the experiment to the TV show data.

Synergies between the variationist and speaker assessment agendas as exemplified in the study of Speaker Design suggest one way in which the justly criticized pitfalls of a-contextual, self-serving experimental research without clear, ulterior implications may be avoided in the future. Certainly, it shows that a constructionist reappraisal of the experimental method as applied in much work on 'language attitudes' does not inevitably force its abandonment. Nor does it signify that automatic preference is to be given to the qualitative elicitation of evaluative practices, such as via discourse analysis (see Hyrkstedt and Kalaja 1998; Liebscher and Dailey-O'Cain 2009 for examples), which may simply serve other purposes of application.

THE ISSUE OF QUANTIFICATION

The supposedly negligent treatment of context effects, as one major point of criticism proffered by social constructionists against traditional scale-based attitude measurement, has been part of a wider complaint about the suppression of potential variability in informants' responses (Hyrkstedt and Kalaja 1998; Potter and Wetherell 1987). Though reference is here mainly made to issues of experimental control (but see my counter-argument above) and researchers' coding impositions, the point easily extends to the very premise of quantitative survey research. Dörnyei (2007: 27) formulates the issue as such:

Because people differ from each other in the way they perceive, interpret, and remember things, their accounts will show considerable variation across individuals. [...] Quantitative researchers regard the sample-related variation as a problem which needs to be fixed. [Their] solution is to take a large enough sample in which the idiosyncratic differences associated with the particular individuals are ironed out by the sample size and therefore the pooled results largely reflect the commonalities that exist in the data. Qualitative researchers, on the other hand, question the value of preparing an overall, average description of a larger group of people because in this way we lose the individual stories. [...] Thus, quantitative researchers follow a 'meaning in general' strategy, whereas qualitative researchers concentrate on an in-depth understanding of the 'meaning in particular'.

To some extent, then, the discussion of the treatment of variability in evaluative activity can be brought back to a fundamental debate about the usefulness of quantitative vs. qualitative research. Certainly, what speaker assessment tasks typically aim for are not differentiated individual accounts, but rather the average mainstream of social meanings associated with variation in language use, by way of the usually concomitant statistical computation of results (often based on comparisons of mean scores or rank differences on the semantic differential scales). In other words, there is no denying that through the analytic routine of quantification and averaging of scores, our experiments render invisible individual informants' rating patterns, reducing them to a common denominator. This seems to once more return us to the issue of the meaningfulness of comprehensive measures; but in fact, such meaningfulness can be argued to be entirely a function of the research goals, rather than a matter of scientific ideology. Recall that in the brief presentation above of my Austrian study, the elicitation of speaker assessments was subservient to the interpretation of conversational contextualization in instances of strategic standard-dialect style-shifting in TV dis-

cussions. Particularly in such a public speaking context, contextualization works by broad consensus and convention: speakers' rhetorical strategies probably rely on the assumption of widely known social associations which the majority, the 'average' addressees/ audience share – in other words, precisely on the types of common denominator of social meanings (we may want to call them 'stereotypes') a quantitative speaker assessment survey is perfectly suited to bringing out. Variability and nuances in judgments are not relevant here – broad-stroked knowledge consensus is.

The question of whether or not to use a broad, quantitative, scale-based survey should thus be one of research intent, and not so much of philosophy. Both approaches have their up- and downsides (for further discussion see Dörnyei 2007; Coolican 2009). And while it is certain that some assumptions of experimental research in (social) psychology are no longer tenable under present-day constructionist epistemology, I hope to also have shown here that a culling of the method from our battery, instead of a reappraisal and adaptation, would be, as the idiom goes, an unfortunate case of 'throwing out the baby with the bathwater'.

CONCLUSION

The bottom line is, then, that the demands of modern social constructionist theorizing and epistemology do not automatically compel the abandonment of our traditional speaker assessment methodology – particularly of using our much-cherished semantic differential scales and the concomitant quantification of findings. What is indeed called for in my opinion, however, is a more rigorous application and specification of what our methodology can do and find, and what it cannot, and of the particular purpose it is to be applied to.

By the same token, I suggest that a reassessment of our terminology is also in order. I have tried to make a case for recasting 'language attitudes' in terms of 'social meanings of linguistic variation'. I believe that this will provide some much-needed impetus for further theorizing and integration of our field, which for too long has been 'overrepresented by one-off studies in widely varying cultures, sociolinguistic conditions, situational and procedural domains' (Giles and Coupland 1991: 49), and has furthermore unnecessarily tried to live up to the standards of traditional social psychological research on 'attitudes'. However, all things considered, I am actually resigned to the fact that the use of the terms

‘language attitudes’ and ‘speaker evaluation’ themselves are far too entrenched (even in my own mind!) to be replaced by any other terminology that might be more accurate to the endeavor, following my line of reasoning. This is probably nothing to worry about – as long as we are clear that what we are saying may be ‘attitude’, but what we are doing is *the measurement of (social) meaning*.

REFERENCES

- Agheyisi, R. and J. A. Fishman. 1970. Language attitude studies: A brief survey of methodological approaches. *Anthropological Linguistics* 12(5): 137–157.
- Albarracín, D., B. T. Johnson and M. P. Zanna (eds.). 2005. *The Handbook of Attitudes*. Mahwah, NJ: Lawrence Erlbaum.
- Bassili, J. N. and R. D. Brown. 2005. Implicit and explicit attitudes: Research, challenges, and theory. In D. Albarracín, B. T. Johnson and M. P. Zanna (eds.) *The Handbook of Attitudes*. Mahwah, NJ: Lawrence Erlbaum. 543–574.
- Bohner, G. and N. Dickel. 2011. Attitudes and attitude change. *Annual Review of Psychology* 62: 391–417.
- Cargile, A. C., H. Giles, E. B. Ryan and J. J. Bradac. 1994. Language attitudes as a social process: A conceptual model and new directions. *Language & Communication* 14(3): 211–236.
- Coolican, H. 2009. *Research Methods and Statistics in Psychology*. 5th ed. London: Hodder Education.
- Coupland, N. 2007. *Style*. New York: Cambridge University Press.
- Dörnyei, Z. 2007. *Research Methods in Applied Linguistics*. Oxford, UK: Oxford University Press.
- Eagly, A. H. and S. Chaiken (eds.). 1993. *The Psychology of Attitudes*. Fort Worth: Harcourt Brace Jovanovich.
- Eagly, A. H. and S. Chaiken. 2005. Attitude research in the 21st century: The current state of knowledge. In D. Albarracín, B. T. Johnson and M. P. Zanna (eds.) *The Handbook of Attitudes*. Mahwah, NJ: Lawrence Erlbaum. 743–767.
- Eckert, P. 2008. Variation and the indexical field. *Journal of Sociolinguistics* 12(4): 453–476.
- Fabrigar, L., T. K. MacDonald and D. T. Wegener. 2005. The structure of attitudes. In D. Albarracín, B. T. Johnson and M. P. Zanna (eds.) *The Handbook of Attitudes*. Mahwah, NJ: Lawrence Erlbaum. 79–124.

- Fishbein, M. and I. Azjen. 1975. *Belief, Attitude, Intention, and Behavior*. Reading, MA: Addison-Wesley.
- Garrett, P. 2005. Attitude measurements/Messung von Einstellungen. In U. Ammon, N. Dittmar, K. Mattheier and P. Trudgill (eds.) *Sociolinguistics / Soziolinguistik: An International Handbook of the Science of Language and Society*. Berlin: Mouton de Gruyter. 1251–1260.
- Garrett, P. 2010. *Attitudes to Language*. Cambridge, UK: Cambridge University Press.
- Gergen, K. J. 2008. On the very idea of social psychology. *Social Psychology Quarterly* 71(4): 331–337.
- Giles, H. and N. Coupland. 1991. *Language: Contexts and Consequences*. Buckingham: Open University Press.
- Greenwald, A. G., D. E. McGhee and J. L. K. Schwartz. 1998. Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology* 74(6): 1464–1480.
- Gumperz, J. J. 1982. *Discourse Strategies*. Cambridge: Cambridge University Press.
- Heise, D. R. 1970. The semantic differential and attitude research. In G. F. Summers (ed.) *Attitude Measurement*. Chicago: Rand McNally. 235–253.
- Himmelfarb, S. 1993. The measurement of attitudes. In A. H. Eagly and S. Chaiken (eds.) *The Psychology of Attitudes*. Fort Worth: Harcourt Brace Jovanovich. 23–87.
- Hyrkstedt, I. and P. Kalaja. 1998. Attitudes toward English and its functions in Finland: A discourse-analytic study. *World Englishes* 17(3): 359–368.
- Kalaja, P. 1997. Language attitudes reconsidered: From cognitive representation to discursive construction. Paper presented at the American Association for Applied Linguistics Annual Conference, Orlando, FL, March 8–11, 1997.
- Kristiansen, T. 2009. The macro-level social meanings of late-modern Danish accents. *Acta Linguistica Hafniensia* 41: 167–192.
- Krosnick, J. A., C. M. Judd and B. Wittenbrink. 2005. The measurement of attitudes. In D. Albarracín, B. T. Johnson and M. P. Zanna (eds.) *The Handbook of Attitudes*. Mahwah, NJ: Lawrence Erlbaum. 21–76.
- Lambert, W. E. 1967. A social psychology of bilingualism. *Journal of Social Issues* 23(2): 91–109.
- Lambert, W. E., R. Hodgson, R. C. Gardner and S. Fillenbaum. 1960. Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology* 60(1): 44–51.

- Liebscher, G. and J. Dailey-O'Cain. 2009. Language attitudes in interaction. *Journal of Sociolinguistics* 13(2): 195–222.
- Likert, R. 1932. A technique for the measurement of attitudes. *Archives of Psychology* 140: 1–55.
- McConahay, J. B. 1985. Modern racism, ambivalence and the Modern Racism Scale. In J. F. Dovidio and S. L. Gaertner (eds.) *Prejudice, Discrimination, and Racism*. San Diego, CA: Academic Press. 91–125.
- Osgood, C. E., G. J. Suci and P. H. Tannenbaum. 1957. *The Measurement of Meaning*. Urbana, IL: University of Illinois Press.
- Potter, J. 1998. Discursive social psychology: From attitudes to evaluative practices. *European Review of Social Psychology* 9(1): 233–266.
- Potter, J. and M. Wetherell. 1987. *Discourse and Social Psychology*. London: Sage.
- Preston, D. R. 2009. The cycle of regard, perception, and production. Paper presented at the workshop 'Production, Perception, Attitude', University of Leuven, April 2–3, 2009.
- Preston, M. S. 1963. Evaluational reactions to English, Canadian French and European French voices. Unpubl. M.A. thesis, McGill University, Montreal.
- Schilling-Estes, N. 2002. Investigating stylistic variation. In J. K. Chambers, P. Trudgill and N. Schilling-Estes (eds.) *The Handbook of Language Variation and Change*. Malden: Blackwell. 375–401.
- Schwarz, N. 2007. Attitude construction: Evaluation in context. *Social Cognition* 25: 638–656.
- Schwarz, N. 2008. Attitude measurement. In W. D. Crano and R. Prislin (eds.) *Attitudes and Attitude Change*. New York: Psychology Press. 41–60.
- Soukup, B. 2009. *Dialect Use as Interaction Strategy*. Vienna: Braumüller.
- Soukup, B. 2010. Matched-guise experimentation as a speech event: A recast and its implications. Paper presented at the Sociolinguistics Symposium 18, University of Southampton, September 1–4, 2010.
- Soukup, B. 2011. Speaker Design in the context of Southern American English: Process models and empirical evidence. *Brno Studies in English* 37(1): 125–138.
- Thurstone, L. L. 1928. Attitudes can be measured. *American Journal of Sociology* 33(4): 529–554.
- Zahn, C. J. and R. Hopper. 1985. Measuring language attitudes: The Speech Evaluation Instrument. *Journal of Language and Social Psychology* 4(2): 113–123.