

Naomi Kamoen. *Positive versus Negative. A cognitive perspective on wording effects for contrastive questions in attitude surveys.* Utrecht: LOT, 2012. XVIII + 175 pages. ISBN 978 94 6093 083 6. EUR 19,36.

To gain access to people's opinions and attitudes, surveys (questionnaires) are generally regarded as "the least cumbersome method" (Dijkstra & Van der Zouwen 1982: 2). However, surveys do not unambiguously measure people's subjective states, as answers can be influenced by a multitude of seemingly irrelevant factors. For example, Rugg (1941) showed that people exhibit an asymmetry in answering *forbid/allow* questions; from a logical perspective, answering *no* to the question 'Do you think the US should *allow* public speeches against democracy?' is the equivalent of answering *yes* to the question 'Do you think the US should *forbid* public speeches against democracy?' However, the percentage of respondents that answered *no* to the *forbid* question was significantly higher than the percentage of respondents that answered *yes* to the *allow* question. The research presented in this dissertation focuses on answering three questions concerning such wording effects:

- 1) Does the *forbid/allow* asymmetry generalize over other contrastive word pairs?
- 2) How can variation in the response effect of adjectival word pairs be explained?
- 3) What are the cognitive mechanisms behind wording effects?

Chapter 2 – The generalizability of wording effects

Meta-analyses show that the *forbid/allow* asymmetry can be generalized, but replications of Rugg's original experiment show a large variety in occurrence, direction and size of effect. In order to determine the generalizability of the effect, Kamoen reports on a series of twelve experiments concerning contrastive word pairs other than *forbid/allow*. The experiments vary with respect to several contextual characteristics, such as survey topic and sampling procedure. Each of the twelve studies examined the effect of 13 contrastive adjectival pairs (*duidelijk (clear)/onduidelijk (unclear)*, *persoonlijk (personal)/onpersoonlijk (impersonal)*, *et cetera*) on mean answers to questions concerning text comprehension and attractiveness. The studies were carried out by twelve MA-students, which resulted in an impressive sample size of 1619 respondents, enabling the measurement of the effects of contrastive word pairs in positive ('X is good. *Yes/No*'), negative ('X is bad. *Yes/No*') and bipolar questions ('X is *good – bad*'). The methods and analyses are sufficiently explained and clear examples are given, as well as details on mean effects, effect sizes and variance. The results show that for each of the word pairs, a significant wording effect was observed in at least one study. The overall tendency is that respondents are more likely to disagree with negative questions, than to agree with positive questions. However, these results differ between studies. For instance, the word pair *logisch (logical)/onlogisch (illogical)* showed a larger effect than *uitnodigend (inviting)/afhoudend (reluctant)*, but the latter effect is far more consistent across studies. The overall analysis shows that for five word pairs, a consistent effect can be found across studies, while the in-study effects of the remaining eight word pairs are likely to be contextual.

Remarks: The appendices at the end of the chapter give insight into the heterogeneity of respondents and survey characteristics. Kamoen argues this heterogeneity to be a "strict test for deciding whether wording effects can be generalized" (p. 20). However, as previous research indicates that context may influence the (quite subtle) wording effects of contrastive questions, certain contextual characteristics could have been factors in the design or variables in the analysis of the results. Instead, choices such as objective wording ('The text is...', 8 studies) vs. subjective wording ('I think the text is...', 4 studies) and the use of 5-point scales (2 studies) vs. 7-point scales (10 studies) were left to the MA-students and were not investigated in detail.

Chapter 3 – Explaining variation between word pairs in wording effects

In this chapter, Kamoen formulates and tests the hypothesis that the variance in wording effects found in chapter 2 might be explained by the semantic distinction between relative and absolute gradable adjectives. Relative gradable adjectival word pairs, such as *tall – short*, are context dependent. A *short* adult is still taller than a *tall* child; there are no fixed endpoints to the scale. In contrast, absolute gradable adjectival word pairs, such as *open – closed*, do have absolute endpoints, irrespective of the object described; a *closed door* possesses a zero degree of *openness*. The crucial difference is that the denial of an endpoint of an absolute adjective entails the affirmation of the opposite pole (*not closed = open*), while this doesn't hold for relative adjectival pairs (*not large ≠ small*). A (large) difference in answers is expected for questions with relative adjectives, but not for questions with absolute adjectives. A prerequisite for testing this hypothesis is that such a distinction can be made irrespective of context. Semantic theory provides a clear heuristic for this: at least one of the endpoints of an absolute adjectival word pair should combine with a degree modifier ('The door is *almost/fully* closed'), while none of the members of a relative adjectival pair should ('The boy is *#almost/#fully* small/large'). As the scale of an adjective is said to be a lexical property, this behaviour should be constant across contexts. 173 participants judged 40 sentences on a 5-point scale (*unacceptable – acceptable*) containing 20 antonym pairs frequently used in surveys (*clear – unclear, easy – difficult, et cetera*). The adjectives were combined with modifiers and placed in different contexts. The results show a correlation between acceptability and type of adjective: an adjective generally combines both with *fully* and *almost*, or fails to combine with either of them. However, this classification only holds within contexts. Between contexts (sentences), modifier-adjective combinations are rated differently. Results show that this variance is due to contextual factors and between-subject effects. Distributed over different contexts, 80% of the estimated acceptability ratings will lie within a 1.29 – 4.31 interval. In other words, there is too much contextual variance to classify adjectival word pairs on lexical grounds.

Remarks: The semantic classification of absolute and relative adjectives fails to meet the requirements to serve as a possible explanation for the variance found in wording effects. Although this is unfortunate from a linguistic point of view, Kamoen presents the method and results in a clear way and the experimental approach is a welcome addition to the field of semantics. Kamoen discusses the implications for survey design and semantic theory and she raises the question whether the absolute/relative property of adjectives is indeed lexical. However, this does not mean that there are no linguistic theories that can do a better job at explaining the variance found. Kamoen briefly suggests some possible explanations from semantic/pragmatic theory, such as non-default adjective interpretations (Bolinger 1972) and scope of modifiers (Kennedy & McNally 2005), but the chapter would have benefited from a more extensive discussion of results in light of linguistic theory.

Chapter 4 – An examination of the cognitive processes in answering contrastive questions

This chapter discusses the question in which of the four cognitive processes of the Tourangeau model (Tourangeau et al. 2000) the wording effects take place. These processes are 1) interpreting the question, 2) retrieving attitudinal information from long-term memory, 3) rendering judgment and 4) fitting judgment to response options. Cognitive activities 1-3 are grouped in the 'comprehension-retrieval stage', while activity 4 forms the 'mapping stage'. Locating the wording effects in one of these stages is important for all research involving surveys. If wording effects arise in the first stage, respondents activate a different judgment about the object in question, leading to the measurement of a different attitude. This would pose a problem for the validity of contrastive questions. Conversely, if the effect arises in the mapping stage, the same underlying attitude is measured for positive and negative questions, while they are fitted differently onto the answering options. This would suggest

that both positive and negative questions are valid measurements of the same attitude. (This does not mean that there are no differences in answers on positive and negative questions.) Experimental studies suggest that the wording effect takes place in the mapping stage (the ‘mapping account’). However, semantic and psycholinguistic literature on negation suggests that negative terms take longer to process than positive terms, predicting an effect in the comprehension-retrieval stage (the ‘cognitive load account’). To decide on this issue, Kamoen has carried out an eye-tracking experiment in which 56 participants answered 90 survey questions. Again, the expected wording effect was observed: respondents were more likely to disagree with negative questions, than to agree with positive questions. This result is ‘only’ an intermediate step in this experiment, but it is noteworthy that Kamoen replicates the wording effect in all of her experiments. Four processing measures were collected: 1) first-pass reading time for question and 2) for answers, 3) remaining fixation time for question and 4) for answers. As the first-pass reading time for the question necessarily reflects the comprehension-retrieval process, the cognitive load account predicts a processing time difference in this stage for contrastive questions. Kamoen finds interesting differences in answering processes for single-reading trials and re-reading trials, but no significant effect of wording on processing time was found. This means that there is no reason to assume a wording effect on either the comprehension-retrieval or the mapping stage.

Remarks: Kamoen gives various reasons why the results speak against the cognitive load account, but not all arguments are equally convincing. For instance, the comprehension-retrieval and mapping division of cognitive processes seems too general to warrant a definitive conclusion. Not finding an effect in either stage does, as Kamoen concludes, possibly form an argument *against* the cognitive load account. This however does not mean that it automatically provides an argument *for* the mapping account. While a more frequent occurrence of re-reading for negative questions and answers was found, this too does not serve as direct evidence for the mapping account, as it could also reflect a difficulty in comprehension. Surprisingly, Kamoen concludes this chapter with the remark that “the results of the current study increase the likelihood that only the mapping process is affected by the choice for a positive or a negative wording of a survey question” (p. 85). This seems not entirely warranted given the research results, but in chapter 5, this statement is repeated even more emphatically (p. 92). Some other general remarks are in order for this chapter. First, 56 people participated in this study, while only half received eye-tracking measurement (p. 66). There is no reason mentioned for this choice, nor are its consequences discussed. Second, the first glance at the response options does not imply a completed comprehension-retrieval stage. Kamoen accounts for this by arguing that the questions were easy and the respondents were skilled (p. 71), but this seems too strong an assumption with respect to the high-precision eye-tracking measurements. Third, in contrast to the first experiment, almost all participants were female university students, which makes it hard to generalize results.

Chapter 5 – A further inspection of the cognitive processes in answering contrastive questions

The measures used in the previous experiment were, according to Kamoen, “very coarse-grained” (p. 92). A final experiment was conducted in order to draw more precise conclusions on the cognitive load vs. mapping debate. Measurements were collected in line with the four-stage Tourangeau model. In the previous experiment, the ordering of the subject, evaluative term (ET), indicator of opinion and attitude object (AO) varied, resulting in residual variance. This experiment has a 2x2 factorial design, which accounts for possible interactions between question polarity and the order of AO (e.g. *animal testing*) and ET (e.g. *unacceptable*). This way, effects on eye movement can be related more directly to cognitive processes. For instance, when the AO receives longer fixation for negative terms only when it follows the ET, it is likely a spill-over effect of evaluative term comprehension. If the AO

receives longer fixation for negative terms regardless of AO-ET order, it is likely that the retrieval of information or the formation of an opinion is also affected, which implies differences in validity of contrastive questions. 122 university students participated in the eye-tracking experiment. The results show that both in single and re-reading trials, the evaluative term receives longer initial fixation when it is negative. For single-reading trials, the response options are read longer for negative questions than for positive questions and for re-reading trials, the evaluative term receives longer fixation when worded negatively. Kamoen concludes that using more fine-grained eye-tracking measures, results show that the comprehension-retrieval stage is affected by the choice of wording. This seems to support the cognitive load account, but Kamoen is right to point out that her results are not decisive; negative terms have lower overall frequencies in language use, which could be an alternative explanation for the effect found. Moreover, the effect seems to be very local: the results show that only the comprehension of the evaluative term itself is affected, which suggests that positive and negative questions are equally valid in measuring attitudes. Kamoen sometimes seems opposed to the cognitive load theory, but the reader must be aware to interpret this only in relation to explaining the wording effects; cognitive load *theory* is supported by the results, but as an *account*, it does not explain wording effects. Kamoen concludes in favour of the mapping account, because for positive questions the attitude object is re-read more frequently, while for negative questions the evaluative term and questions are re-read more often. The chapter ends with the conclusion that negative questions are slightly more difficult to comprehend than positive questions, but that wording effects probably arise when translating opinions into the response options given. Therefore *yes* to a positive question is not the same as *no* to a negative question.

Remarks: This chapter is a clear report of an experiment that extends the results reported on in the previous chapter. Kamoen does seem to have a preference for the mapping account from the start of chapter 4, while evidence for this account does not arise until the end of chapter 5. This may confuse the reader, as this preference sometimes conflicts with results found. To prevent this, the distinction between cognitive load and mapping as *effects* or *phenomena* and as explanatory *accounts* could have been made clearer. A final remark is that in this chapter too, re-reading is seen as an indication of mapping processes only, which ultimately rests on the assumptions that the questions were not difficult and that participants were skilled in survey answering.

Chapter 6 – Synthesis, general remarks

The last chapter offers a synthesis of the research results and a discussion of the implications for survey practice, semantics and methodology. Kamoen suggests that the fact that answers are given relative to evaluative terms forms an explanation why respondents are more likely to disagree with negative questions, than to agree with positive questions. A problem concerning this conclusion is that it does not really explain the *why* of the effect, but rather the *when* or *where*. The relation between evaluative term and response options does not explain the wording effect on the level that is expected when reading this dissertation. Chapter 3 gears towards a semantic explanation of the effect found, but unfortunately, the theory proved incapable of explaining the effects. This is not Kamoen's fault, but it does relate to a more general comment on this dissertation. The book's chapters are published articles, which this makes the reading experience somewhat repetitive. Moreover, there is no real transition between chapters. Chapter 3 follows up on chapter 2 nicely, because it examines a possible explanation for variance found in the preceding chapter. However, chapter 4 and 5 seem to address another question: "*when or where* does the effect arise?" instead of "*why* does it arise?" These questions are obviously connected, but they are not the same. Arguments for this shift are only briefly mentioned, but the reader is sometimes left wondering in what direction the book is heading, as there is no reflection on the implications of not finding an explanatory account in chapter 3. The focus on

methodology and survey research is a quality of this dissertation, but it may have stood in the way of embedding the research more firmly in linguistic theory. For instance, the difference between an implicit negative adjective (*unhappy*) and an explicitly negated adjective (*not happy*), seems relevant to the findings, but is only very shortly addressed in chapter 5. Here the book could have benefited from syntactic, semantic and pragmatic research on negation, such as Horn and Kato (2000), and from the emerging field of experimental pragmatics (Noveck & Sperber 2004). Furthermore, Kamoen only briefly mentions that *positive* and *negative* in her experiments are understood in evaluative terms in a footnote on page 46. This does not account for the fact that in some word pairs ‘pure’ negation is involved (*desirable* vs *undesirable*), while other words are paired with evaluative counterparts (*fascinating* vs *boring*). Argumentative semantics could have provided useful suggestions, such as distinguishing between *argumentational* and *evaluative* orientation (Verhagen 2000, 2005). If the research had been embedded more in linguistic theory, the discussions of results could have been richer, and the results might have had a wider impact.

To summarize, this book gives a detailed view of the effect of contrastive question wording, and the focus on methodology gives insight into a series of interesting and well-designed experiments. While the elaborate statistical analyses and discussions may deter some linguists, the experiments, results and conclusions are in fact valuable to both survey research and linguistics.

Alex Reuneker
a.reuneker@umail.leidenuniv.nl

Bibliography

- Bolinger, D. (1972). *Degree words*. The Hague: Mouton.
- Dijkstra, W., & Van der Zouwen, J. (1982). *Response behavior in the survey-interview*. London: Academic Press.
- Horn, L., & Kato, Y. (Eds.). (2000). *Negation and Polarity: Syntactic and Semantic Perspectives*. Oxford: Oxford University Press.
- Kennedy, C., & McNally, L. (2005). Scale structure, degree modification, and the semantics of gradable predicates. *Language*, 81, 345–381.
- Noveck, I., & Sperber, D. (Eds.). (2004). *Experimental Pragmatics*. Basingstoke: Palgrave Macmillan.
- Rugg, D. (1941). Experiments in wording questions. *Public Opinion Quarterly*, 5, 91–92.
- Tourangeau, R., Rips, L., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge: Cambridge University Press.
- Verhagen, A. (2000). “The girl that promised to become something”: An exploration into diachronic subjectification in Dutch. In F. S. Thomas & J. P. Snapper (Eds.), *The Berkeley Conference on Dutch Linguistics 1997: the Dutch Language at the Millennium*. (pp. 197–208). Lanham, MD: University Press of America.
- Verhagen, A. (2005). *Constructions of intersubjectivity. Discourse, syntax, and cognition*. Oxford: Oxford University Press.