

Visually mediated valence effects in dialogue: an explorative study

Richard Andersson¹

Rasmus Bååth¹

Sverker Sikström²

¹Lund University Cognitive Science

²Department of Psychology, Lund University

Abstract

This is an exploratory study investigating potential effects of emotional valence in images and their influence on conversation in the presence of the images. We used latent-semantic analysis to generalize valence ratings of Swedish words to a corpus of spoken conversations. Each utterance in the conversation was given a valence rating, which represented how emotionally positive or emotionally negative the utterance was. We found no effects that indicate that valenced images have an effect on conversations. However, we find that valenced images in general, and positive images in particular, were considered more helpful by the participants who engaged in the conversations. Additionally, we find no results that interlocutors align over time in their use of valenced language.

Keywords: valence, emotions, communication, dialogue, alignment, latent semantic analysis

Richard Andersson
Lund University Cognitive Science
Kungshuset, Lundagård
SE – 222 22, Lund
Sweden
richard.andersson@humlab.lu.se

Visual information, valence and dialogue

Communication makes use of shared knowledge between the speakers, and this knowledge can be anything from a history of previous exchanges, shared cultural background, and having access to the same perceptual input from our environment. This is usually simply summed up as two speakers having a “common ground” (Clark, 1996; Stalnaker, 2002). Having common ground in terms of a shared visual workspace has been found improve (here: speed up) collaborative problem-solving processes by allowing the language to be used in reduced form, reduce clarification questions and so forth (e.g. Kraut, Fussell & Siegel, 2003). This effect is easy to intuitively understand, as it simply means that because we both see the same information, and we know that we both see the same information, we can take it for granted that we both have this information accessible to be referred to without overly specifying descriptions.

Previous studies in language and vision have mainly focused on reference resolution and how integrating visual information can help this process (e.g. Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy, 1995; Knoeferle, Crocker, Scheepers & Pickering, 2005; Glenberg & Robertson, 1999). However, we believe there are also other, neglected, venues of scientific interest.

One particular interesting research question is whether the available visual information also changes the communicated content in a dialogue on a more general level apart from making references to the material. For example, we may be biased by the present information and take the perspective of the visual information as a presupposition in our communication. More concretely, if we e.g. discuss the famine in developing countries and at the same time have access to images of the effects of famine – is our conversation content influenced? One idea is that the visual information activates concepts and idea relating to famine, and this in turn influence what subtopics are raised in our conversation. In this case it would likely activate concepts such as suffering, poverty, war as a cause, and other concepts easily activated by images. If, on the other hand, the present images instead displayed a map of developing countries, charts of development indexes, et c., would the same concepts be activated and used in the conversation? Probably not. Unless the speakers already have a strong viewpoint and perspective, it seems likely that they would use the information available (alternatively, the information would use them). This idea is typically the concern of media studies concerning “framing effects” it is under-studied there as well (c.f. Rodriguez & Dimitrova, 2011). That humans can be influenced by subtle cues is uncontroversial and many astounding effects have been found, e.g. that participants in a mock language experiment using words related to old age (“elderly”, “wrinkle”, “Florida”), internalized these meanings and were found to walk to the next

Andersson, R., Bååth, R. & Sikström, S. Lund University Cognitive Studies, 151, 2012.

experiment room more slowly than participants who were primed with neutral words (Bargh, Chen & Burrows, 1996).

Concerning the domain in initial example, images of famine are strongly linked to the processing of emotional content (valence) in an image. One idea is that emotional content in visual information will also influence the content of a conversation such that words tend to become more emotional, i.e. the shared discourse of the discussants are rich in emotional words. Research concerning the processing of emotional stimuli have found some interesting differences between valent stimuli depending on if it has positive (such as flower, babies and puppies) or negative content (such as death, suffering and betrayal). Although the results are often mixed, in general negative stimuli seem to be better at capturing and holding the attention of participants (Blanchette, 2006; Fox, Russo, Bowles, & Dutton, 2001). The emotional content of an image also seem to modulate the perceptual span of a viewer, with negative material narrowing the attention span, so that viewers may not attend to and get information about other parts of a visual scene (Fernandes, Koji, Dixon & Aquino, 2011). However, positive information seems to be processed faster (Unkelbach, Fiedler, Bayer, Stegmüller & Danner, 2008).

To translate these effects into changes in a dialogue is not straight-forward, but one attempt is to say that negative visual information will hold the attention of the participants, disrupting their conversation productivity and possibly forcing them to discuss the visual information, resulting in the use of “negative” words. Positive visual information, on the other hand, are processed faster and will result in increased productivity (compared to the conversation of the negative material) for the speakers and the use of more “positive” words. However, another intuition is that negative images may not be the worst conversation starters as they may still force some kind of reaction due to being overly positive or negative about a topic or providing a reference point to distance your self against. A completely neutral and bland image is perhaps the least facilitating for conversation productivity. Given that we know that images accompanying a conversation can be viewed as helpful by speakers and shown to improve the communication (Andersson, Holsanova & Holmqvist, submitted), we can also expect that positive images are viewed as more helpful by the speakers and that their communicative success (*ibid.*) may be facilitated.

That visual information can prime and influence speakers in a dialogue is not incompatible with current psycholinguistic models of dialogue. For example, in the interactive alignment model of dialogue processing (Pickering & Garrod, 2004), speakers become primed by each others use of words, concepts and syntactic structures, so they in turn use the same forms themselves. The authors of this model argue that this priming spreads across all levels, leading to interlocutors becoming more similar in their language and thinking over time, i.e. they achieve “alignment”. For our purposes, we can view present visual information as another source of

Andersson, R., Bååth, R. & Sikström, S. Lund University Cognitive Studies, 151, 2012.

priming, which will initiate a cycle of priming among its viewers which will align over time to the associations activated by the visual information. For this purpose, we expect to see that participants align over time in their use of emotionally charged language and that this process is faster when given access to visual information.

Latent semantic analysis

As we now have developed some intuitions concerning what effects to expect in conversations under the influence of emotional visual information, the next challenge is how to analyze conversations and their emotional content. One approach is to simply have coders rank all utterances in conversations according to the valence of the utterances. Another approach, which we ultimately settled for, is to approximate the meaning of the utterances by constructing a high-dimensional space based on the co-occurrence of words in different linguistic contexts such as documents or paragraphs. In this space, words that co-occur will be located closer to each other, which also gives them their meaning. For example, the geometric word “circle” occurs with other geometric words, so a word like “radius” would be located fairly near to “circle” and share some meaning, whereas a word like “guitar” would be located much further away and represent something other than geometrical concepts. After a dimensionality reduction of this space, we end up with a matrix where the estimated similarity of a word to another can be calculated as the cosine of the angle between the two word vectors in the matrix. So, words which are highly similar will have a cosine closer to 1 and words which are highly dissimilar will have a cosine closer to 0. This model even has the property of determining the similarity of words which never occur together, the so-called *latent* meaning. The technique for creating and using these models is known as *latent semantic analysis* (LSA; Deerwester, Dumais, Furnas & Landauer, 1990).

The validity of LSA for building models of meaning have been verified in several studies. For example, Landauer & Dumais (1997) created an LSA model that performed equally to successful applicants on a TOEFL test, given a training corpus representative of a college freshman’s complete text exposure since birth. The vocabulary growth rate of the model also mimicked the rate of American children. Huettig, Quinlan, McDonald, & Altmann (2006) used an LSA model to extract the semantic relationships between words to show that the spoken words activated semantic information which in turn influenced eye-movements to semantically related objects in a scene. The results confirmed previous findings regarding eye movements to semantically related objects in language experiments and thus showed that the LSA model captured the expected semantic similarity between words.

The LSA model was used in the current study to generalize a number of rated valence words to all words in the model. The valence values of the words was then used to calculate the average valence in a particular utterance or conversation.

Hypotheses

This study is an explorative study and so we have not concentrated our analysis to one or two particular hypotheses. Rather, we wish to gain insight into what mechanisms are promising venues of further research and see what effects are realistic to expect in less restricted conversation settings.

Images influencing word content

Our first set of hypotheses simply concerned the question whether the valence of the image contaminated the valence of the conversation, such that positive images led to a more positive conversation, and vice versa.

H1: A positive correlation between rated image valence and LSA-estimated average word valence in an utterance.

Image valence and conversation productivity

Our second set of hypotheses test whether valent images stimulate the productivity in a conversation. This hypothesis comes in two variants, where one tests whether positive image valence stimulates conversation productivity due to the positive-valence bias for processing information. The other hypothesis tests if any valence, i.e. whether an image is positive or negative but not neutral, stimulates the conversation productivity. An intuition is that even negative images can facilitate production by being provocative conversation-starters. These effect could possibly be manifested as both an effect on the number of words per utterance, but also on the number of utterances per conversation. Both of these levels was explored.

H2a: A positive correlation between the rated image valence and the number of words per utterance.

H2b: A positive correlation between the absolute image valence and the number of words per utterance.

H2c: A positive correlation between the rated image valence and the number of utterances per conversation.

H2d: A positive correlation between the rated absolute image valence and the number of utterances per conversation.

Valent images and their perceived helpfulness

Our third set of hypotheses concerned the subjective experience of the helpfulness of the images and whether participants found valent images to be more helpful in this conversation task.

Andersson, R., Bååth, R. & Sikström, S. Lund University Cognitive Studies, 151, 2012.

H3a: A positive correlation between rated image valence and rated image helpfulness.

H3b: A positive correlation between rated absolute image valence and rated image helpfulness.

Image valence and communicative success

Our fourth set of hypotheses dealt with the communicative success of the interlocutors. This success was estimated by the predictive success of interlocutors in estimating their conversation partner's opinion (see Andersson *et al*, submitted, for the reasoning behind this measure).

H4a: A negative correlation between rated image valence and absolute prediction error.

H4b: A negative correlation between rated absolute image valence and absolute prediction error.

Visual context and valence alignment

Our fifth and final set of hypotheses concerned the idea of valence alignment over time. The hypotheses reflected the questions if there exist such a thing as valence alignment and whether this can be modulated by visual context information.

H5a: A negative correlation between utterance valence differences between interlocutors and time approximated by utterances index.

H5b: A negative correlation between utterance valence differences between interlocutors and the interaction of time and the presence of visual context.

H5c: A negative correlation between utterance valence differences between interlocutors and the interaction of time and the rated valence of the visual context.

Method

This study analyzes data recorded during the course of another experiment (Andersson *et al*, submitted), but a brief summary of the method will be given here.

Participants

The participants consisted of 48 participant pairs. All participants spoke Swedish at native level and were compensated for their participation in the experiment.

Materials & Design

The conversation topics were based on 48 statements concerning current events in society. These statements were matched with scenes relevant to the topic. For example, the conversation statement: “Overall, the train traffic in Sweden works satisfactorily”, was matched to an image of a very busy train station. It should be noted that the images were selected “pseudo-randomly” by the first author from the internet to illustrate the topic statements. The conversations were evenly split through a mixed-lists design in two different conditions: either with the matched image, or without the matched image. Only eight statement–image pairs, determined by counter-balancing, were used per participant pair.

After each conversation topic, participants rated how well they and their partner have contributed to a successful conversation. They also estimated their engagement level for this particular conversation topic. All questions were answered privately.

After the conversation part of the experiment, participants answered questions regarding their and the other's opinion, how helpful they believed the different images were, how relevant they were and their perceived valence in conjunction with the topic. Other questions, not analyzed at the time of writing, outside the scope of this paper followed after these questions. All questions were estimated using a 7-point Likert scale (1–7 or -3–+3).

The responses concerning the other's (predicted) opinion were corrected using each participants (true) opinion rating. This prediction error formed the dependent variable concerning communicative success where lower error means better communication (as proposed in Andersson *et al*, submitted).

The LSA model was based on 100,000 articles taken from the 100 largest Swedish newspapers (including web-based ones) in 2007. The high-dimensional model consisted of 20,000 contexts (each a span of 30 words) and the 15,000 most common words in the source material. The data was iteratively (500 times) reduced by means of singular value decomposition so in the end each word was represented by a 100 dimensional vector. In this form, words with similar meaning are represented with similar vectors. The model predicted the valence of the words using a multiple linear regression with the 100 semantic dimensions as predictors. These valence rankings were then validated by a leave-one-out cross-validation procedure, where a word with known valence was estimated using all other words in the model. This procedure was performed iteratively until all words had been predicted. This resulted in a significant correlation between the known and the estimated valences ($r = .419$; $p < .001$). The two most positively valent utterances as well as the two most negatively valent utterances are shown in Table 1.

Procedure

Participants were briefly introduced and asked to sit down at opposite sides of a table. Next to them was a 24 inch computer monitor which displayed all written instructions, and later the statement–image pairs. Participants discussed each conversation statement for five minutes, with the statement (and possibly the image, depending on condition) present at the computer screen at all times. After each conversation, they were allowed a short break as they filled out a questionnaire. After all eight topics had been discussed, each participant was asked to take a seat in front of a computer to answer the computerized questionnaires. Finally, participants were debriefed and given their compensation.

Table 1: Example utterances with LSA-estimated valence scores

Valence	Swedish	English translation (approx.)
9.3906	... på nåt vis, även om det är som du säger intressant tycker jag.	“... in some way, even if it is, like you say, interesting, I think.”
9.3131	En glad kaffearbetare får ju mig att tänka på rättvisemärkt, och då kommer rättvisefrågor in i [xxx], in i, he-he, i det hela också.	“A happy coffee worker makes me think about fair trade certification, and then fairness issues come into play too.”
3.0451	... från personer med vapen.	“... from people with weapons.”
2.9878	... som har vapen och äh ...	“... who have weapons and uh...”

Analysis

The statistical analysis was done by means of mixed-effects regression models using the lme4 package (Bates, Maechler & Bolker, 2011; R Development Core Team, 2011; Baayen, Davidson & Bates, 2008). The regressions were specified for either a Gaussian- or Poisson-distributed data using their default link functions (identity and log, respectively). Poisson models used an observation-level intercept to compensate for overdispersion. The models used random intercepts to adjust according to the inherent variability of participants and items, where some participants or pairs e.g. produce more or less valent utterances compared to others, and where some items help generate more or less valent utterances than others. Random slopes were also included if they were warranted according to a log-likelihood ratio test (LRT). Random slopes were not warranted unless otherwise reported. Hypothesis tests were carried out by different procedures, depending on what is appropriate and technically possible with the current implementations of the statistical software packages (MCMC simulation for Gaussian data with no random slopes,

Andersson, R., Bååth, R. & Sikström, S. Lund University Cognitive Studies, 151, 2012.

LRT for gaussian data with random slopes, and Wald test for Poisson-distributed data). Reported p-values are uncorrected for the multiple comparisons. Concerning the robustness of the results, the significance level equivalent to an α of .05 with $n = 12$ hypothesis tests is, according to the

$$\frac{1}{12}$$
Šidák correction (Šidák, 1967): $1 - (1 - \alpha)^{1/12} = 0.0043$. This assumes that the tests are independent, which they are not, so the true significance level is somewhere between .0043 and .05 (closer to the former rather than the latter).

Results

The results are summarized in Table 2 and described more thoroughly in the following subsections. The participants produced a total of 34575 utterances. The average interlocutor produced 47 utterances (SD = 19) which resulted in an average of 458 words (SD = 148).

Images influencing word content (H1)

The regression model indicates that conversations, as the mean valence of an utterance, does not change due to being in the presence of images with varying valence. However, the data set of ratings for the image valence was incomplete due to the design of the original experiment where only half of the shown conversation topics were accompanied by an image. To investigate whether this could be a power issue that could be solved, we imputed the missing values using a random multiple imputation technique, and then re-ran the models. The non-significance of the results remained.

Image valence and conversation productivity (H2)

We found no effect that images influenced conversation productivity in the form of number of words per utterance – no matter if we coded images as ranging from negative to positive, or just coding their absolute valence as neutral to valent. Furthermore, we tested the hypotheses that valent images could stimulate productivity in the form of the number of utterances per conversation. We found no such effect, neither for raw nor absolute valence ratings of the images. As in H1, we re-ran the models after having imputed the missing values. Also as in H1, the non-significance of the results remained.

Valent images and their perceived helpfulness (H3)

Random slopes were not warranted according to an LRT, but we included them in order to handle the heteroskedasticity of the data. The analyses indicated that images were perceived as helpful both when using the raw valence ratings and the absolute ratings. This means that as images become more positive they are experienced as more helpful for the conversation. It also

Andersson, R., Bååth, R. & Sikström, S. Lund University Cognitive Studies, 151, 2012.

means that as images deviate more from the neutral middle-point on the valence scales, whether that be in the positive or in the negative direction, the images are experienced as more helpful. Whereas the first result is significant but not obviously reliable considering the multiple comparisons problem, the second results is reliable even after correcting for multiple hypothesis tests. As the ratings of image helpfulness are right-skewed, we transformed the rating to a binary variable with ratings ≤ 0 coded as “not helpful” and > 0 as “helpful”. Re-running the analyses with this binary outcome variable confirmed the initial significant results.

Image valence and communicative success (H4)

According to LRTs, random slopes were warranted for both hypotheses. We found no support for the hypotheses that more valent images should stimulate conversation so that their communicative quality increased (according to the prediction-based measure). It made no difference whether we considered the raw or absolute valence ratings of the images. Imputation of missing values did not change the results.

Visual context and valence alignment (H5)

Time was coded as the log index of an utterance. To reduce the presence of back-channelling or other utterances that were not part of a proper turn, we focused the analyses to only include utterances with 4 words or more. This operation did not change the results however. We found no effect of alignment at all and it did not matter whether we included the presence of images or the valence of the images as predictors.

Table 2: Statistical results

Hypothesis	Intercept	Estimate	Standard error	Test statistic ($z/t/\chi^2$)	P-value
H1	6.865	0.006	0.008	t = 0.70	0.469
H2a	1.736	-0.002	0.011	t = -0.23	0.853
H2b	1.741	-0.004	0.016	t = -0.29	0.792
H2c	87.810	-0.753	1.944	t = 0.026	0.980
H2d	89.803	-2.548	2.707	t = -0.941	0.468
H3a	0.908	0.104	0.059	$\chi^2 = 4.250$	0.024*
H3b	0.476	0.390	0.087	t = 4.468	0.001***
H4a	-1.810	-0.034	0.528	z = -0.064	0.949
H4b	-1.844	0.519	0.489	z = 1.062	0.288
H5a	-0.318	-0.029	0.110	z = 0.271	0.787
H5b	-0.302	0.105	0.221	z = 0.474	0.635
H5c	-0.327	0.151	0.170	z = 0.885	0.376

Discussion

This study aimed to investigate a number of effects relating to the emotional charge of utterances and the effect of conversation-present images with varying degrees of emotional charge.

Of all our hypotheses, only the hypotheses concerning the helpfulness of the valent images were significant. However, this effect did not carry over to the measure of communicative success, which has previously (Andersson *et al*, submitted) been found to be influenced by the rated helpfulness of the images. It is unlikely that this is simply a power issue, as our imputation attempts hardly changes the results. Perhaps it is simply the case that participants do not have the insight they think they have. They say that valent images are helpful, but that does not mean that the images are truly helpful or valent, or that they use the images during conversations in the way they believe they use them.

We found no effect at all that participants become influenced in the emotional content of their conversations due to the presence of valent images. This is surprising, partly due to there being a conflated variable in the design. Imagine that we have participants that simply mention what they see on the screen. If the screen then contains the negatively valent image of a shot

reindeer, then words such as “shot”, “dead” and perhaps “killed” would have occurred. This would in turn have affected the utterance valence in the analysis in the negative direction and worked toward the alternative hypothesis. However, as we see no such effect, then either participants are not affected at all, they choose not to show that they are affected, or that the effect is small and perhaps located just in the first one or two utterances after image onset. A plausible explanation of the lack of effects is perhaps that the images used in the experiment were part of an study where explicitly valent images were not desired. The conversation-accompanying images are mostly neutral images such as picture of train stations, power plants, academic environments, and so on. The valence range of the images was perhaps simply too narrow to elicit our wanted effects, especially considering the range in images of picture data bases such as the International Affective Picture System (IAPS: Lang, Bradley & Cuthbert, 2008). However, the limited valence range also work to our advantage. The majority of visual input we process each day is of very limited emotional strength. Thus, had we found positive results using the IAPS, then it would not necessarily mean that the effects would be likely to occur in everyday language use. As we tested fairly neutral images directly, we cannot say that these valence effects are unlikely to occur for strongly valenced images, but we have shown that it is less likely that the effects would occur in an everyday situation. In one view, it may be a relief to find that participants are not that affected by conversation-present images, which in a real-world situation could mean e.g. biased news footage.

One interesting null effect which is outside the challenges with the pictorial stimuli, is that interlocutors do not align over time with regard to the valence of their utterances. In other words, it is not so that participants become more emotional in their language over time as the discussion heats up, nor is it so that speakers have separate styles, where one is the positive optimist and the other is the negative cynic, and that they agree on a style of time. The interactive-alignment model does predict that participants align on all levels.

Andersson, R., Bååth, R. & Sikström, S. Lund University Cognitive Studies, 151, 2012.

Acknowledgements

RA & RB gratefully acknowledges support from the Linnaeus environment Thinking in Time: Cognition, Communication and Learning (CCL), Swedish Research Council grant number 349-2007-8695.

Andersson, R., Bååth, R. & Sikström, S. Lund University Cognitive Studies, 151, 2012.

References

Andersson, R., Holsanova, J. & Holmqvist, K. (submitted). Opinion prediction as communicative success, and the role of visual information in conversations.

Baayen, R.H., Davidson, D.J. and Bates, D.M. (2008) Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59, 390-412.

Bates, D., Maechler, M. & Bolker, B. (2011). lme4: linear mixed-effects models using S4 classes. (Version 0.999375-39) [computer software]. Available from: <http://CRAN.R-project.org/package=lme4>

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, 71 (2), 230-244. doi:10.1037/0022-3514.71.2.230

Blanchette, I. (2006). Snakes, spiders, guns, and syringes: How specific are evolutionary constraints on the detection of threatening stimuli? *Quarterly Journal of Experimental Psychology*, 59, 1484 –1504.

Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41(6), 391-407.

Fernandes, M. A., Koji, S., Dixon, M. J. & Aquino, J. M. (2011). Changing the focus of attention: The interacting effect of valence and arousal. *Visual Cognition*, 19 (9), 1191–1211.

Fox, E., Russo, R., Bowles, R., & Dutton, K. (2001). Do threatening stimuli draw or hold visual attention in subclinical anxiety? *Journal of Experimental Psychology: General*, 130 (4), 681–700.

Glenberg, A. M. & Robertson, D. A. (1999). Indexical understanding of instructions. *Discourse Processes*, 28(1), 1-26.

- Andersson, R., Bååth, R. & Sikström, S. Lund University Cognitive Studies, 151, 2012.
- Huetting, F., Quinlan, P. T., McDonald, S. A., & Altmann, G. T. M. (2006). Models of high-dimensional semantic space predict language-mediated eye movements in the visual world. *Acta Psychologica*, 121(1), 65-80. doi:10.1016/j.actpsy.2005.06.002.
- Knoeferle, P., Crocker, M.W., Scheepers, C., & Pickering, M.J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: evidence from eye-movements in depicted events. *Cognition*, 95, 95-127.
- Kraut, R. E., Fussell, S. R., & Siegel, J. (2003). Visual information as a conversational resource in collaborative physical tasks. *Human-Computer Interaction*, 18, 13-49.
- Landauer, T. K. and Dumais, S. T. (1997) A solution to Plato's problem: the Latent Semantic Analysis theory of acquisition, induction and representation of knowledge. (html) *Psychological Review*, 104(2) , 211-240.
- Lang, P.J., Bradley, M.M., & Cuthbert, B.N. (2008). International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8. University of Florida, Gainesville, FL.
- Pickering M.J. & Garrod S. (2004). Towards a mechanistic Psychology of dialogue. *Behavioral and Brain Sciences*, 27, 169-226.
- R Development Core Team (2011). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
- Rodriguez, L., Dimitrova, D. V. (2011). The levels of visual framing. *Journal of Visual Literacy*, 30 (1), 48–65.
- Šidák, Z. (1967). Rectangular confidence regions for the means of multivariate normal distributions. *Journal of the American Statistical Association*, 62, 626–633.
- Stalnaker, R. (2002). Common Ground. *Linguistics and Philosophy*, 25, 5-6, 701-721.

Andersson, R., Bååth, R. & Sikström, S. Lund University Cognitive Studies, 151, 2012.

Tanenhaus, M. K., Spivey-Knowlton, M.J., Eberhard, K. M. & Sedivy, J. E. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268, 1632–1634.

Unkelbach, C., Fiedler, K., Bayer, M., Stegmüller, M., & Danner, D. (2008). Why positive information is processed faster: The density hypothesis. *Journal of Personality and Social Psychology*, 95(1), 36-49. doi:10.1037/0022-3514.95.1.36