"Searching for Happiness" or "Full of Joy"? Source Domain Activation Matters

MEYLYSA TSENG, YIRAN HU, WEN-WEI HAN, BENJAMIN BERGEN
University of Hawai'i at Manoa

0. Introduction

> Happiness does not depend on outward things, but on the way we see them.
> — Leo Tolstoy

The differences among words with similar meanings are notoriously hard to pin down. What differentiates a *foe* from an *enemy*? When is a large wooded area a *forest* and when a *wood*? Fine distinctions in lexical semantics, apparent with concrete language like this, are even harder to address for language of a more abstract nature. How are we to distinguish, if it is indeed possible to do so, between *freedom* and *liberty*; *ideas* and *thoughts*; or *constancy* and *permanence*? Semantic subtleties like these torture second language learners, feed the imagination of poets, and furnish language mavens with editorial careers. Our aim in the current work is to investigate in some depth what contributes to the selection of one word over another extremely similar word in language use. In particular, the question at hand is whether lexical choices are influenced by differences in the metaphorical patterns that the words in play participate in. The investigation of how metaphorical patterns influence lexical choice is useful not just as an end unto itself; it also provides a key to the role of metaphorical grounding in the mental representation of abstract words more broadly.

The literature on near-synonyms is small but varied. It shares a focus on distinguishing among similar words on the basis of distributional usage patterns. However, the lexical representations constructed by such distributional pattern differences are used for diverse purposes, such as linguistic theory building, language education and natural language processing technology. One typical example is a study of the differences in near-synonymous verbs in Mandarin Chinese. Chief, et al. (1998) looked at the near-synonym pair *fangbian* 'be convenient' and *bianli* 'be convenient'. Through a corpus investigation, they found that each verb has different distributional patterns when nominalized, negated, and used as a verbal predicate, nominal modifier, verbal modifier,

transitive and intransitive. In addition, they also differ in the frequency with which they take sentential or verbal objects vs. complex nominal objects. Chief, et al. (1998) used their findings within the framework of lexical semantics to posit that *fangbian* profiles the whole proposition event and *bianli* profiles the beneficial role of the event.

Another approach to differentiating between near-synonyms is to propose contrastive semantic features. This was done in Tsai, et al. (1998) to further develop lexical semantic theory and in Inkpen and Hirst (2004) to further advance natural language generation systems. Specifically, Inkpen and Hirst (2004) extract from entries of near-synonym dictionaries, such as *Choose the Right Word* (Hayakawa 1994), semantic features pertaining to level of formality, denotation and attitude. Feature specifications thus include detailed information of different types, for instance that a word such as *fib*, for example, is informal, whereas the near-synonym *prevarication* is formal. These near-synonyms, with their feature specifications, are stored in a lexicon which is accessed, along with another database constructed from distributional probabilities from corpora, when the program needs to produce a word in a certain context. This context either has the features specified by the user, or is determined by the program analyzing the surrounding text.

Such contrastive features, though helpful in machine translation and language processing applications,[1] only tell part of the story about the meanings underlying similar words. For one, they are insufficient to support real human understanding because they lack grounding in a human(-like) conceptual system and body (Barsalou 1999, Zwaan 1999, Glenberg and Robertson 2000). Without access to grounded conceptual knowledge about how words are used, knowing even detailed facts about the linguistic contexts in which *above* and *over* are used doesn't tell a language user much about what they mean or how their meanings are different from *under* and *below*.

A second drawback of contrastive features is that they fail to take into consideration important advances in metaphorical theory. The importance of metaphor in determining the meanings of abstract language is crucial, since it has been shown that the collocates of a given target word are often metaphorical (Stefanowitsch 2004). This should not be particularly surprising since Lakoff and Johnson (1980) demonstrated in their seminal work that language about concrete, perceptually and motorically grounded domains, like containment, can be applied equally well to abstract domains, like emotions. In language use these abstract domains are conceptualized in terms of these concrete domains (as in *full of anger*). The field of enquiry known as Conceptual Metaphor Theory, built upon this observation, has provided evidence that these metaphorical patterns of

[1] For other uses of corpus-based lexicons see Fillmore and Baker (2001) for using FrameNet (located at http://framenet.icsi.berkeley.edu/papers/FNcrime.pdf) in natural language processing technology.

language are indicative of deeper metaphorical patterns of thought. Namely, just as abstract concepts are described in terms of concrete source domains, so they are understood in terms of those domains. Historically, the evidence for this tenet of Conceptual Metaphor Theory has been predominantly linguistic in nature, consisting of 1) lists of linguistic expressions used metaphorically according to similar metaphorical mappings, and 2) demonstrations that novel metaphorical language takes new meanings along the same lines defined by those purported conceptual relations across domains, though other evidence, like transfer of inference, is also cited (Lakoff and Johnson 1980, Lakoff 1993).

While abstract concepts, and thus the words that denote them, appear to be constructed metaphorically, it is not universally accepted that metaphorical linguistic expressions are built on top of conceptual mappings across domains. One of the strongest criticisms of Conceptual Metaphor Theory's "understanding-as" account of abstract language is the scarcity of experimental evidence that the metaphors used to express abstract concepts trigger source domain conceptual structures (Murphy 1996, 1997). Accounts based on linguistic patterns alone could be construed as merely "an imaginative cataloguing of etymological relics" (Boroditsky 2000:6), and must be complemented with independent empirical evidence.

This call for convergent evidence has not gone unanswered. Psycholinguistic research by McGlone and Harding (1998), Gentner, et al. (2002) and Boroditsky (2000, 2001) has provided evidence that language from a source domain primes both concepts and language selection in the target domain. Boroditsky (2000, 2001) also showed that language about the source domain activates source domain conceptual structures, which then activate similar target domain conceptual structures. Boroditsky's (2000) third experiment is especially informative, since it suggests that though source domain language primes target domain reasoning, target domain language does not prime source domain reasoning.

Specifically, Boroditsky (2000) created an online experiment where subjects were primed with either two spatial or two temporal true or false questions, followed by a third target true or false question, which was either spatial or temporal. Spatial questions for prime and target stimuli consisted of a picture and a true or false statement describing the picture, such as "M is in front of me" or "X is in front of M." Temporal questions were true or false statements such as "In March, May is ahead of us" or "March comes before May." Her independent variables were the type of prime and target (spatial or temporal) and whether the metaphor of the primes was consistent with the metaphor of the target question (consistency vs. inconsistency). The dependent variable was the time it took subjects to respond to the critical question.

Boroditsky found that reaction times were longer for inconsistent target statements only when spatial primes occurred before spatial or temporal target questions, and when temporal primes occurred before temporal target questions. Thus, subjects use the spatial domain to think about space and time, and they use the time domain to think about time. However, subjects do not use the time

domain to think about space, providing evidence for a unidirectional mapping from the source to the target domain.

The approach of activating a metaphorical source domain and testing the effects on target domain behavior can equally well be applied to near-synonyms. In the current study, this is precisely the tack taken. Instead of studying effects on a target domain reasoning task, we investigate whether the activation of a metaphorical source domain leads to the choice of an abstract word which tends to be described metaphorically in terms of that same source domain. The two near-synonyms we investigated were the words *joy* and *happiness*. These words, while quite similar in meaning, have been shown through corpus research (Stefanowitsch 2004) to exhibit different tendencies with respect to the metaphors used to describe them. *Happiness* has a tendency to be described as a thing to be searched for and acquired, while *joy* tends to be talked about as though it were a liquid that fills the body. Applying a modified version of the source domain priming method described above, then, we hypothesized that if language users rely on source domains to make lexical choices about abstract language, then when asked to make a choice between the two words *joy* and *happiness*, subjects who are in the process of searching should be more likely to choose the compatible *happiness*, while subjects who are drinking liquids should be more likely to select the word *joy*.

1. Experiment 1: SEARCHING and DRINKING

The goal of this first study was to test whether experiencing a concrete source domain primes the choice of a word that pertains to an abstract domain that is metaphorically spoken about in terms of that same source domain. We relied on distributions for metaphors pertaining to the words *happiness* and *joy* provided by Stefanowitsch (2004). His study of a corpus of German and American English newspapers investigated the distribution of metaphorical language used with these words and their German translation equivalents.

Critically, some of the metaphors he catalogued are strongly biased toward one word. For the purposes of our study, HAPPINESS IS SEARCHED FOR occurs much more frequently than JOY IS SEARCHED FOR. In contrast, the metaphor JOY FILLS A CONTAINER is more frequent overall than HAPPINESS FILLS A CONTAINER. These source domains, SEARCHING and FILLING A CONTAINER, were strongly associated with their respective target words. The ratio of *happiness* to *joy* in the SEARCHING metaphor is roughly 7:1 (157:22), and the one for *joy* to *happiness* in the CONTAINER metaphor is 3:1 (35:12). Even with the larger ratio occurring for the SEARCHING metaphor, both words occur significantly more often in their respective metaphors, thus giving us two distinct source domains, one more strongly associated to the target domain of *happiness* and one more strongly associated to the target domain of *joy*.

1.1. Method

With this source and target domain information, a simple questionnaire was designed with a 12.67cm x 15.21cm black and white picture of a person smiling. A scaled-down version of the questionnaire can found in Appendix A. The person in the picture could be described as experiencing either happiness or joy. Underneath the picture in 28 point font was the question "What emotion is this person experiencing?" Underneath this question were the two choices: "A) JOY B) HAPPINESS." Approximately half the subjects saw this ordering of choices, while the other half were shown the questionnaire with "A) HAPPINESS B) JOY."

There were four female experimenters. They looked in libraries, bookstores, grocery stores, coffee shops, restaurants, dormitories and bars for people searching for something or drinking something. In order to find subjects engaging uniquely in searching or drinking behavior, the experimenters did not approach people who were searching for beverages or in the beverage aisle of the grocery store, nor people who were in the process of buying a beverage, or who had just bought a beverage.[2]

When the experimenters found a potential subject, they approached the person and asked for permission to ask three questions. The experimenter then told the willing subject to look at the questionnaire and choose the first answer that came to mind. The experimenter then lifted the questionnaire up to show it to the subject, in the process also covering her own face. In addition, the experimenter was sure to hold the paper either at the top or with both hands in the same position on both sides. After the subject read the question and responded, the experimenter asked the remaining two questions - the subject's age and whether they were a native speaker of English. If there was more than one subject together in a group, subjects were instructed to decide an answer in their head first, and then when all subjects had an answer, they were asked one at a time to tell the experimenter what that answer was.

A total of 164 subjects participated in the experiment, 84 in the survey with choice A as HAPPINESS and choice B as JOY, which will be referred to as the H/J survey, and 80 in the survey with choice A as JOY and choice B as HAPPINESS, the J/H survey. Ages ranged from 18 to 70 years old, with an average age of 34. In the H/J survey there were 13 women and 25 men in the drinking condition and 26 women and 20 men in the searching condition. In the J/H survey there were 10 women and 32 men in the drinking condition and 17 women and 21 men in the searching condition.

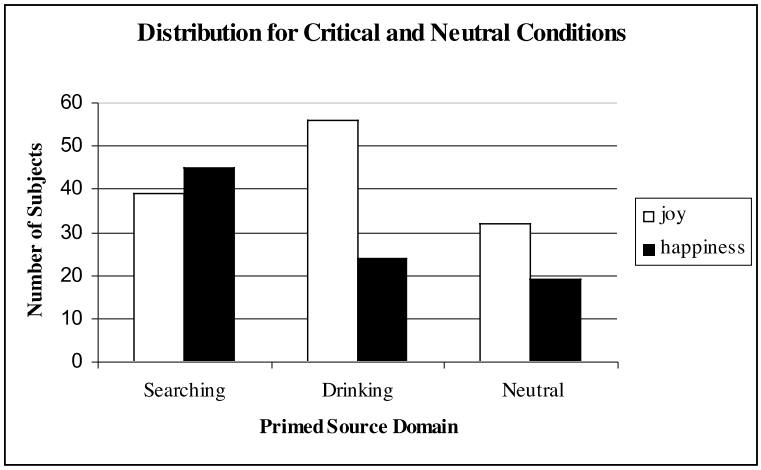
1.2. Results

The overall distribution of *happiness* and *joy* responses across the two surveys shows the expected effect, as seen in (1) below. There were comparatively more

[2] Stefanowitsch (2004) also found that *happiness* occurred more often in the HAPPINESS IS TRANSFER metaphor, which was the other potential confound we were avoiding.

happiness responses when subjects were engaged in searching behavior, 54%, and more *joy* responses when they were engaged in drinking, 70% ($p < 0.01$, two-tailed Fisher's exact test). Looking more closely at the two survey forms individually, the distribution in the J/H survey shows a significant effect in the expected direction ($p < 0.01$, two-tailed Fisher's exact test). The H/J survey, though in the expected direction, only approaches significance ($p = 0.128$, two-tailed Fisher's exact test). A further analysis of effects by gender showed that male and female subjects have results nearing significance in the expected direction (males $p = 0.075$, females $p = 0.198$, two-tailed Fisher's exact test), but due to insufficient numbers these effects were not significant on their own.

(1) Source domain activation priming results for *happiness*, *joy* and neutral



1.3. Discussion

The results show a significant overall effect consistent with the prediction derived from Conceptual Metaphor Theory, namely that lexical choice of nearly synonymous abstract words should be influenced by the activation of their corresponding source domains. This finding suggests that the semantic representations of such abstract words are closely tied to the source domains they are metaphorically understood in terms of. However, this result does not tell us exactly what mechanism is responsible for this priming effect – whether the different lexical choices in the searching and drinking conditions resulted from facilitation of the consistent word, inhibition of the inconsistent word, or both. This concern is addressed in a second experiment, described below.

2. Experiment 2: Neutral Case

The results from Experiment 1 showed that subjects were more likely to categorize a depicted person as experiencing *happiness* when they themselves were in

the act of searching, but *joy* when in the act of drinking. In order to determine whether this tendency was due to facilitation of a consistent word, inhibition of an inconsistent word, or both, we presented the same material to another group of subjects in a neutral condition. This neutral condition was students in classroom situations, since subjects in such a context were unlikely to be in the process of searching or drinking. In order to ensure this was the case, the experiment was conducted well after students sat down, to avoid the effects of searching for a seat. In addition, after giving their responses, subjects were asked to report whether they had brought a drink with them to class, and those who had were excluded from the analysis. Once again, the neutral condition served to provide a baseline frequency for *happiness* and *joy* choices, to allow a tighter analysis of the source of the effect observed in Experiment 1.

2.1. Method

The experiment was administered by four experimenters working individually, three of whom were female and one male. The experimenter first asked the class if they would be willing to participate in a brief survey. After the class agreed, the experimenter told them to look at the same questionnaire used in experiment 1 and choose an answer. Some classes saw the J/H questionnaire and some saw the H/J questionnaire. Students were instructed not to say the answer out loud. After showing the entire class the survey, with hand and survey placement following the same protocol as in Experiment 1, the experimenter passed out survey sheets which asked them to circle their answer, displayed in the same order (J/H or H/J) as on the survey they had just seen. Though the procedure differed in that subjects provided written, rather than oral responses, there is no obvious way that this would influence responses one way or the other, since the order of the two words on the page was counterbalanced.

A total of 75 students participated in the experiment, with ages ranging from 18 to 54 years old. Subjects were students in classrooms at the University of Hawai'i at Manoa. Only undergraduate classes participated. Four were Introductory Linguistics classes, two were Chinese classes, one was a Japanese class and one was a Japanese Linguistics class. Subjects did not receive any compensation for participating. All subjects were native speakers of English.

2.2. Results

Twenty-four subjects were eliminated since they had either circled neither answer, circled both answers (and may or may not have crossed one of them out), were not native English speakers, or indicated that they had brought a drink to class. In the H/J survey, there were 12 women and 14 men and in the J/H survey there were 9 women and 16 men. Overall, 19 people chose *happiness* and 32 people chose *joy* (Figure 1). The two surveys had virtually identical results: in the H/J survey 10 chose *happiness* and 16 chose *joy*, and in the J/H survey 9 chose *happiness* and 16 chose *joy*. In comparing these results with those of the drinking and searching conditions, it was found that the searching condition had signifi-

cantly more *happiness* responses than in the neutral condition ($p = 0.001$, Binomial) but that the drinking condition did not have significantly more *joy* responses than the neutral condition ($p = 0.923$, Binomial).[3] Breaking the results down by survey, for both surveys, in the searching condition, *happiness* responses were significantly more frequent than in the neutral condition (J/H survey $p < 0.05$, H/J survey $p = 0.01$, Binomial). Finally, there was also a gender difference, with only female subjects choosing *happiness* significantly more than *joy* in the searching condition (females $p < 0.001$, males $p = 0.330$, Binomial).

2.3. Discussion

The results of the experiment in the neutral condition show that, when sitting in a classroom, students are almost twice as likely to choose *joy* over *happiness*. This is nearly the same distribution obtained for subjects in the drinking condition. Thus, the results from the neutral condition were significantly different from those in which subjects were searching, but not different from situations in which they were drinking. Assuming that the neutral condition gives a good measure of how likely subjects were to pick each of the words without any particular contextual bias, engaging in searching behavior causes more *happiness* responses, whereas drinking does not have any significant effect on lexical choice.

3. General Discussion

It is not surprising that the searching effect would be much stronger than the drinking effect, since in the linguistic distribution of the SEARCHED FOR metaphor, *happiness* outnumbers *joy* seven to one, whereas in the FILLING A CONTAINER metaphor, *joy* outnumbers *happiness* only by three to one. Another possible explanation for the drinking condition having weaker effects is that the stimulus picture was already biased toward *joy* responses. Thus, there was less room for additional *joy* responses.

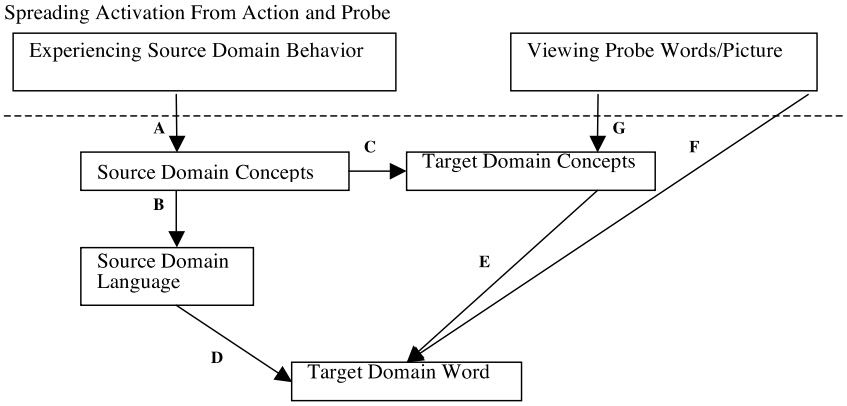
Of course, this interpretation of the data presumes that the neutral condition used in this study was completely neutral. Possible confounds include the fact that students could be in a "searching for knowledge" frame of mind or a "filling themselves with knowledge" frame of mind. The first mindset would cause more happiness responses, while the second would cause more *joy* responses. Thus, it is possible that results were biased toward *joy* due to factors unrelated to the type of picture used.

The results reported above indicate that experiencing a source domain can activate a word that is canonically metaphorically described using that source domain. What sort of mechanism could account for this effect? The direction of activation flow can be modeled from two main sources: 1) performing the source domain action and 2) seeing the picture and probe word. This is diagrammed in figure (2).

[3] Due to the presence of a neutral condition, we analyzed the results with a Binomial instead of Fisher's Exact.

Thus, source domain behavior leads down pathway A, activating the source domain. The source domain then spreads activation down pathway B, causing linguistic processing of source domain language, and pathway C, which spreads activation to all connected target domains. Source domain language then spreads activation to target domain language, as indicated by pathway D, while target domain activation also leads to activation of target domain language, as indicated by pathway E. Thus, lexical access is influenced by both source domain language and target domain conceptual structures.

(2) Processing Routes of Lexical Access



When the subject sees the probe words/picture and reads the critical question, activation is sent down pathways F and G, activating target domain concepts and target domain language. The active target domain then activates target domain language as well, following pathway E. This route shows how the words *happiness* and *joy*, in particular, were accessed for lexical processing and production.

Note that only routes originating from A (experiencing the source domain) predict priming from the source domain. Routes originating from G and F (viewing probe words/picture) have no pathway to the source domain, since pathway C between the source and target domains is unidirectional. Evidence for this, again, was shown by Boroditsky (2000), who found that target domain language did not prime source domain language. Thus, it is hypothesized that routes beginning with A (experiencing the source domain) and leading to E (target domain word) are the processing routes used by subjects to discriminate between the near-synonyms *happiness* and *joy*.

3. Conclusion

This experiment tested whether engaging in source domain behavior would activate metaphorically linked target domains, and thereby influence lexical choice between near-synonyms. The source domains of searching and drinking

provided a fruitful testing ground because they provide contrastive metaphorical language for target domain words *joy* and *happiness* (Stefanowitsch 2004). We found that in choosing between near-synonyms to describe a particular picture, subjects who were searching were more likely to choose *happiness* than *joy*, and subjects who were drinking or in the neutral condition were more likely to choose *joy* than *happiness*. One explanation for the stronger effect of searching is that the dominance of the metaphorical pattern HAPPINESS IS SEARCHED FOR is much more pronounced than the dominance of the pattern JOY FILLS A CONTAINER (i.e 7:1 vs. 3:1). In order to further test this strength of metaphorical pattern account, other target words with highly unique and strongly correlated source domains need to be tested.

In conclusion, this study demonstrated that engaging in source domain action primes target domain language, influencing lexical choice between near-synonyms. Thus, semantic theories as well as natural language production technology would be well served to take steps to further investigate the cognitive effects of metaphorical patterns. Specifically, source domain information from the environment could be a necessary element of the human language processor, making concrete some of the previously undefined nuances that allow us to choose between nearly synonymous abstract language.

References

Barsalou, Lawrence. 1999. Perceptual symbol systems. *Behavioral and Brain Sciences* 22:577-609.

Boroditsky, Lera. 2000. Metaphoric structuring: understanding time through spatial metaphors. *Cognition* 75(1):1-28.

Boroditsky, Lera. 2001. Does language influence thought?: Mandarin and English speakers' conceptions of time. *Cognitive Psychology* 43:1-22.

Chief, Lian-Cheng, Chu-Ren Huang, Keh-Jiann Chen, Mei-Chih Tsai and Lili Chang. 1998. What can near synonyms tell us. In Miriam Butt and Tracy Holloway King, eds., *Proceedings of LFG98 Conference*. CSLI Publications.

Fillmore, Charles J. and Collin F. Baker. 2001. Frame semantics for text understanding. Paper presented at WordNet and Other Lexical Resources Workshop.

Gentner, Dedre, Matsumi Imai and Lera Boroditsky. 2002. As time goes by: Evidence for two systems in processing space and time metaphors. *Language and Cognitive Processes* 17(5):537-565.

Glenberg, Arthur M. and David A. Robertson. 2000. Symbol grounding and meaning: A comparison of high dimensional and embodied theories of meaning. *Journal of Memory and Language* 43:379-401.

Hayakawa, S. I., ed. 1994. *Choose the Right Word*. Second Edition, revised by Eugene Ehrlich. HarperCollins Publishers.

Inkpen, Diana Zaiu and Graeme Hirst. 2004. Near-synonym choice in natural language generation. In Nicolas Nicolov, Kalina Bontcheva, Galia Angelova,

and Ruslan Mitkov, eds., *Recent Advances in Natural Language Processing III*. Amsterdam: John Benjamins Publishing Company.

Lakoff, George. 1993. The contemporary theory of metaphor. In A. Ortony, ed., *Metaphor and Thought (2nd Edition)*. Cambridge: Cambridge University Press.

Lakoff, George and Mark Johnson. 1980. *Metaphors We Live By*. Chicago: University of Chicago Press.

McGlone, Matthew S. and Jennifer L. Harding. 1998. Back (or forward?) to the future: the role of perspective in temporal language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 24(5):1211-1223.

Murphy, Gregory L. 1996. On metaphoric representation. *Cognition* 60:173-204.

Murphy, Gregory L. 1997. Reasons to doubt the present evidence for metaphoric representation. *Cognition* 62:99-108.

Stefanowitsch, Anatol. 2004. HAPPINESS in English and German: a metaphorical-pattern analysis. In M. Achard and S. Kemmer, eds., *Language, Culture and Mind*. Stanford: CSLI Publications.

Tsai, Mei-Chih, C.-R. Huang, K.-J. Chen, and K. Ahrens. 1998. Towards a representation of verbal semantics: an approach based on near synonyms. *Computational Linguistics and Chinese Language Processing* 3(1):61-74.

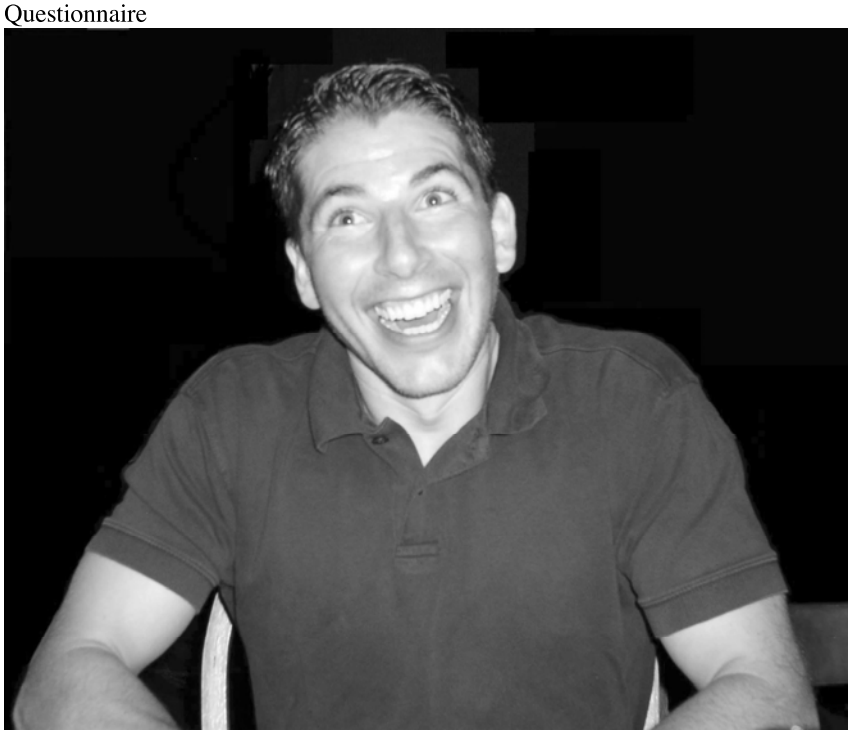
Zwaan, Rolf A. 1999. Embodied cognition, perceptual symbols, and situation models. *Discourse Processes* 28:81-88.

Department of Linguistics
1890 East-West Rd, #569
University of Hawai'i at Manoa, Honolulu
Hawai'i, HI 96822

meylysa@hawaii.edu

Appendix A

Questionnaire



What emotion is this person experiencing?

A) JOY **B) HAPPINESS**