LANGUAGE AND EMOTION

Language and emotion: Putting words into feelings and feelings into words

Kristen A. Lindquist,¹ Maria Gendron,² & Ajay B. Satpute³

1. University of North Carolina at Chapel Hill
2. Northeastern University
3. Pomona College

To appear in The Handbook of Emotions, 4th edition

Address correspondence to:
Kristen Lindquist
University of North Carolina
Chapel Hill, NC 27599
kristen.lindquist@unc.edu
According to the poet Seamus Heaney, putting feelings into words is "a search for images and symbols adequate to our predicament" (Feeling into Words, Preoccupations). Of course, a poet would know better than anyone why it is important to understand the link between emotion and language. Humans have the unique capacity to experience complex, nuanced, emotions. Humans also have the unique challenge of communicating those experiences to one another with language. To date, much research has investigated how our emotional experiences get translated into language; this research is important for domains ranging from the arts to therapy to cross-cultural communication. Yet what Heaney didn’t acknowledge is that the symbols we know might also shape how we experience our “predicament” in the first place. That is, language might not just translate feelings into words, but might help shape the nature of those feelings to begin with.

Throughout this chapter, we review the various traditions that have investigated relationships between language and emotion. These traditions stem from different areas of research (e.g., psychology, neuroscience, linguistics, anthropology) and often make different assumptions about the nature of the relationship between language and emotion. We begin our chapter by first offering a few definitions of what we mean by “language” and “emotion,” in the first place. We next discuss accounts of the relationship between language and emotion. We first introduce an account that explicitly hypothesizes that language helps to constitute emotions—the psychological constructionist model of emotion. Next, we review evidence from the emotion regulation literature that assumes that language can modulate emotions after the fact (e.g. by virtue of “reappraisal” or “affect labeling”
as described below). Finally, we discuss a literature on the emotion lexicon, which typically (with some exceptions from linguistics) focuses on how emotional experiences get translated into words for the sole means of communication. We close our chapter by suggesting that a psychological constructionist approach can unite findings from across these seemingly diverse domains, by describing how words help shape the emotions that people experience and perceive, by proposing the ultimate mechanism by which words help regulate emotions, and by explaining cultural variation in how emotions get put into words.

Definitions

Before we begin, it bears mention what we mean by the terms “language” and “emotion” throughout this chapter. We use “language” to refer exclusively to the words that people use to describe emotional states (such as “anger,” “disgust,” “fear,” “joy,” “contentment,” “pride,” “schadenfreude,” “amae,” etc.). More specifically, we are referring to what linguists call the “semantic” aspects of language. It is beyond the scope of this chapter to discuss the role of syntax or metaphor in emotion, or how the affective meaning of words impacts their understanding or use (although these are interesting areas of study in their own right). Nor do we discuss words that name other categories (e.g., “mother,” “murder”) that might themselves have emotional connotations (for a brief comment on the implications of these types of words in emotion, see Lindquist, MacCormack & Shablack, under review).

We also have a very particular meaning in mind when we use the term “emotion” throughout this chapter. As it turns out, there is no single agreed upon
scientific definition of the term “emotion.” In keeping with the psychological constructionist approach we take in our own research, we thus use the term “emotion” to refer to what are sometimes called “discrete emotions” in the psychology literature—psychological states that are experienced as coordinated patterns of physiology, behavior, and thoughts that occur within certain types of situations, and which are described with certain emotion category words (e.g., in English, “anger,” “disgust,” “fear,” “happiness,” “sadness,” etc.). We differentiate “emotions” from “affect,” which consists of basic feelings from the core of the body (for this reason, it is sometimes called “core affect”; Barrett, 2006b; Barrett & Bliss-Moreau, 2009; Russell, 2003). Affect is the representation of the body’s ever-changing internal state (from the smooth muscles, skeletal muscles, peripheral nervous system, and neurochemical/hormonal system) and is often described as a homeostatic barometer that allows an organism to understand whether objects in the world are good for it, bad for it, approachable or avoidable (Barrett & Bliss-Moreau, 2009).

Throughout, we will differentiate between experiences of emotion (or affect), which we identify as feelings in one’s own body (e.g., a feeling of anger; a feeling of unpleasantness) and perceptions of emotion (or affect), which we define as inferring emotional feelings in another based on their face, voice, body, behavior and so on (e.g., seeing someone else as angry; seeing someone else as feeling unpleasant). We turn now to the psychological constructionist approach to emotion, which explicitly hypothesizes that the words someone knows for emotion shapes how they make
meaning of affect, turning those affective states into emotion experiences and perceptions.

**Putting words into feelings: Language and the psychological construction of emotion**

Psychological constructionist views are a family of psychological and neuroscience models that predict a constitutive role of language in emotions. According to *psychological constructionist views*, emotions are experienced when affective states are made meaningful as specific instances of the emotion categories that exist in a given culture. Emotions are thus considered the resulting products, or constructions, of more basic psychological “elements” (Barrett, 2006; Clore & Ortony, 2013; Cunningham, Dunfield, & Stillman, 2013; Lindquist, 2013; Russell, 2003). According to our particular psychological constructionist approach, called the Conceptual Act Theory (CAT; cf., Barrett, 2006b, 2009, 2012, 2014), these more basic elements are representations of sensations from inside the body (affect), representations of sensations from outside the body (e.g., vision, audition), conceptual knowledge about the emotion categories experienced in one’s culture, and executive attention (for reviews see Barrett & Bar, 2009; Barrett & Bliss-Moreau, 2009; Barrett & Satpute, 2013; Lindquist & Barrett, 2012; Lindquist, Wager, Kober, Bliss-Moreau, & Barrett, 2012). A person experiences an emotion when conceptual knowledge is used to make meaning of her core affect (in experiences of emotion), or another person’s facial muscle movements (in perceptions of emotion), tailoring them to the context at hand. Importantly, rather than viewing emotions as
physical types of categories that all individuals are both with, psychological constructionist views conceive of the emotions that are constructed out of more basic elements as nominal kind categories that exist only by nature of fact that a group of people agrees about their features (e.g., a culture agrees that “anger” is an emotion that occurs in contexts when norms are violated, when people scowl, and when vasodilation and an increased heart rate occurs, whereas “fear” is an emotion that occurs in contexts when uncertain events occur, when people gasp, when vasoconstriction and an increased heart rate occurs) (for a discussion see Barrett, 2012; Lindquist, Gendron, Oosterwijk, & Barrett, 2013).

Although all psychological constructionist views agree that emotions are nominal kind categories constructed out of more basic elements, not all views see a role for language in this process. The CAT is unique in that it explicitly describes a role for language in the construction of emotion, insofar as language supports the acquisition and use of concept knowledge in humans (cf., Lindquist, MacCormack, & Shablack, under review; Lindquist, Satpute, & Gendron, in press). The CAT predicts that language plays an especial role in emotion because emotion categories (e.g., “anger,” “fear,” “disgust”) are abstract concepts with highly variable instances (e.g., Barrett, 2006a; Cacioppo, Berntson, Larsen, Poehlmann, & Ito, 2000; Mauss & Robinson, 2009; Wilson-Mendenhall, Barrett, & Barsalou, 2013; Wilson-Mendenhall, Barrett, Simmons, & Barsalou, 2011) (for a discussion, see Lindquist et al. under review). Language helps humans represent all category knowledge (see Lupyan, 2012), but may be especially important to representing abstract categories that do not have strong perceptual regularities in the world (see Barsalou & Wiemer-
Hastings, 2005; Lindquist et al., under review; Lupyan, 2012). In the case of abstract categories, words are a form of “glue” that holds the concept together (for a discussion see Barrett & Lindquist, 2008; Lindquist et al., under review). The word “anger” is thus thought to be in part constitutive of an angry feeling because it supports the category knowledge that is brought online to make meaning of a rapidly beating heart, high blood pressure, and unpleasantness when a person’s trust is violated, or to make meaning of a calmly beating heart, decreased blood pressure and pleasantness when a person enacts revenge. This does not mean that a person needs to speak the word “anger,” or even think it when making meaning of an affective state. Instead, the idea is that “anger” groups a population of instances in a person’s conceptual knowledge (involving representations of sensations from the body, behaviors, and the context) that are all conceived of as members of the same category despite what otherwise might be large differences between them. For instance, within the behavioral domain, punching, running away, kicking, smiling, crying and scowling can all occur in an instance of “anger.” Without that word to bind them, and based on their perceptual similarities alone (what it looks like or feels like to punch, run, kick, smile, etc.), these instances might otherwise belong to different categories and be experienced as such (see Lindquist, MacCormack & Shablack, under review for a discussion). During experiences or perceptions of affect, labels connected to concepts shape the conceptual information that is brought to bear when making meaning of sensations in the environment (a “label feedback hypothesis”; cf., Lupyan, 2012; for a discussion of how the label feedback
hypothesis applies to emotion, see Lindquist, MacCormack & Shablack, under review).

Mounting evidence from social cognitive, neuropsychological, cross-cultural, and neuroimaging studies is consistent with the psychological constructionist view that language helps constitute emotion. It is beyond the scope of this chapter to discuss all of this evidence since we have done so extensively elsewhere (Barrett, 2009; Barrett, Lindquist, & Gendron, 2007; Barrett, Mesquita, Ochsner, & Gross, 2007; Lindquist, 2013; Lindquist & Gendron, 2013; Lindquist et al., under review; Lindquist et al., in press). However, we discuss several illustrative examples of how emotion words help to create the types of discrete emotional experiences and perceptions that people feel in their own bodies or see in the bodies of others.

Perhaps most notably, evidence shows that impairing participants’ access to the meaning of words impairs their ability to perceive emotion on faces, even in tasks that do not require language. This finding suggests that language reaches relatively “deep” into visual perception to help construct a perception that someone else is “angry” or “sad.” For instance, temporarily impairing access to an emotion concept by having participants repeat a word (e.g., “anger”) out loud thirty times (an procedure called “semantic satiation”) makes participants slower and less accurate to perceptually match two facial expressions (e.g., judge that two faces are both angry; Lindquist, Barrett, Bliss-Moreau, & Russell, 2006).

We recently demonstrated that when language is permanently impaired due to a neurodegenerative disease called semantic dementia, so too is emotion perception. Semantic dementia occurs following neurodegeneration in the left
anterior temporal lobe (ATL) of the brain, an important hub in a network supporting the representation of concept knowledge (Patterson, Nestor, & Rogers, 2007). Patients were asked to freely sort pictures of facial expressions (associated with the categories “anger,” “disgust,” “fear,” “sadness,” “happiness,” and “neutral”) into meaningful piles—a task that didn’t require linguistic responses. Whereas healthy control participants sorted the facial expressions into six or more distinct piles for the six emotion categories represented in the set, patients instead sorted faces into 3-4 piles representing positive, negative, and neutral feelings (Lindquist, Gendron, Barrett, & Dickerson, 2014). These findings converge with earlier evidence from a patient, LEW, who could no longer access and use words following a stroke (Roberson, Davidoff, & Braisby, 1999). LEW had difficulty sorting facial expressions of emotion into consistent piles and created several different configurations of piles across multiple testing instances.

Just as impairing access to emotion words impairs emotion perception, there is evidence that making emotion words more accessible alters how emotional facial portrayals of emotion, and even one’s own body feelings, are made meaningful as discrete emotional perceptions and experiences. In a recent experiment (Nook, Lindquist, & Zaki, in press), we assessed whether the presence of words in a repetition-priming paradigm task made participants faster and more sensitive to perceive emotions on faces than did the mere presence of other faces. On each trial, participants viewed a rapidly presented facial emotion expression (the cue stimulus) followed by a second emotional stimulus (the target stimulus) and indicated whether or not the emotion categories reflected in the cue and target
stimuli matched or not. On some trials, the target was a second facial expression 
(*face-face trials*) and on other trials, it was an emotion category word (*face-word 
trials*). Consistent with the hypothesis that increased accessibility to emotion words 
facilitates discrete emotion perceptions, we found that pairing emotional faces (e.g., 
a pouting face) with emotion labels (e.g., “sad”) increased individuals’ speed and 
sensitivity in perceiving emotions on faces. By contrast, participants’ less sensitive 
judgments on face-face trials were driven by the similarity of facial features 
between cue and target (e.g., the presence of a furrowed brow), which did not 
necessarily differ between different emotion categories. In fact, the findings on face- 
face trials suggest that participants were not likely to spontaneously perceive faces 
in terms of discrete emotion because they were focusing on more basic perceptual 
similarities between faces. Participants did not tend to perceive a face in terms of a 
discrete emotion category unless a discrete emotion category word was present to 
facilitate this judgment.

There is also growing, yet still preliminary, evidence that increasing 
accessibility to emotion words while someone experiences a state of unpleasant 
affect alters the particular discrete emotion that he or she experiences. To test this 
hypothesis in a behavioral experiment (*Lindquist & Barrett, 2008a*), we first 
increased participants’ access to the concepts of fear v. anger (i.e., “primed” those 
concepts) by asking participants to write a story about a fearful or angry character. 
In a control condition that did not prime any emotion concepts, participants wrote a 
neutral story about characters discussing a neutral event. We next separately 
manipulated whether participants felt unpleasant and highly activated v. neutral by
asking them to listen to unpleasant v. neutral music and relive past experiences in which they felt unpleasant v. neutral. We hypothesized that making the concept fear more accessible prior to making participants feel unpleasant would cause them to make meaning of their unpleasant state as an instance of fear. To test this hypothesis, we assessed participants’ degree of risk aversion (i.e., the perception that the world is full of danger), which is consistent with how people view the world when they are in a fearful state. As predicted, those participants primed with the word “fear,” who listened to unpleasant music subsequently behaved in a fearful manner (i.e., saw the world as full of danger).

Another recent study demonstrated that the accessibility of emotion concepts influences how feelings manifest as physiological responses during emotion. Participants who were asked to report on their emotions (thereby activating emotion concepts), and who were then berated by an experimenter while completing a challenging mental arithmetic task showed relative increases in total peripheral resistance (Kassam & Mendes, 2013), consistent with physiological responses to perceived threats. On the other hand, participants who did not have emotion concepts activated experienced less total peripheral resistance. Labeling states actually reduced heart rate and cardiac output (Kassam & Mendes, 2013), which are more generally related to feelings of arousal. These findings are consistent with other growing evidence that categorizing a stress response with an emotion word associated with adaptive responding (e.g., excitement) v. an emotion word associated with non-adaptive responding (e.g., fear) alters the resulting
emotional experience, physiology, and behavior (for a review see Jamieson, Mendes, & Nock, 2013).

Another area of research consistent with the psychological constructionist view that language helps constitute emotion comes from research investigating the role of language in the acquisition of new emotion categories. This research is both correlational, tracking the development of language and its relation to emotion perception across early development, and experimental, demonstrating that pairing exemplars with words helps adults acquire novel emotion categories over and above mere experience with category exemplars. The correlational evidence suggests that prior to the development of language, infants are unable to perceive discrete emotions on faces, although they can differentiate pleasant, unpleasant and neutral expressions relatively well (e.g., after habituating to happy faces, five-month-olds look longer at any unpleasant face, whether fearful, angry or sad; Bornstein & Arterberry, 2003). Like infants, two-year-olds, who only know the very simple emotion words “sad” and “happy,” can only reliably differentiate between unpleasant and pleasant facial expressions (e.g., they mistakenly perceive all unpleasant faces as “sad,” even in a task that does not require language). Yet as we mentioned in the first section, as 3- and 4-year-olds begin to learn the words “anger” and “fear,” they can correspondingly differentiate between sad, angry, and fearful facial expressions (for a review see Widen, 2013).

Presumably, words help adults learn the meaning of novel facial expressions in much the same way as they help infants and children over the course of development, and experimental research is consistent with this hypothesis. For
instance, in the absence of words to label distinct faces, adults were unable to perceive novel Chimpanzee facial expressions (e.g., “play,” “scream,” “bared teeth” and “hoot” faces) as distinct from one another. In the first phase of the experiment, adults viewed pictures of unfamiliar Chimpanzee facial muscle movements (e.g., a bared teeth or screaming face) or viewed the faces while also learning to associate them with nonsense words. Participants were later shown images taken from a continuous morphed array of two facial expressions (e.g., ranging from bared teeth to a scream) and were asked to indicate when two faces from the array were similar to one another, and when they were different. Participants who had learned to associate faces with a label displayed “categorical perception”—they were able to perceive a categorical boundary at the midpoint in the morphed array of bared teeth and scream faces—but participants who did not learn to associate faces with a label did not perceive such a categorical distinction (Fugate, Gouzoules, & Barrett, 2010).

Critically, if language helps constitute emotion, then people who speak languages with different emotion concepts should not just communicate emotion differently, but should also perceive and experience it differently. Consistent with this idea, a recent study found that participants who speak different languages see facial expressions differently. The authors used a computer graphics program to measure and then reconstruct East Asian and Western Caucasian participants’ visual representations of facial expressions associated with the categories “happy,” “surprised,” “fearful,” “disgusted,” “angry,” and “sad.” Whereas Western Caucasian participants represented each of the six categories with a distinct set of facial
movements, East Asian participants did not (Jack, Garrod, Yu, Caldara, & Schyns, 2012).

Our own recent work finds that emotion words guide emotion perception, but only when the words match the concepts that a culture regularly uses to make meaning of emotional faces. Individuals from a remote culture, the Himba of Namibia, and the United States were asked to complete an emotional face-sorting task like that completed in the study of semantic dementia patients. Participants completed the task either with or without emotion words as cues (e.g., a cue that a pile should contain “anger”/“okupindika” faces). Without emotion word cues, even US participants did not sort in line with an assumed “universal” model of emotion based on the English emotion concepts “anger,” “disgust,” “fear,” “sadness,” “happiness” and “neutral.” Yet with emotion word cues, US participants sorted more in line with this “universal” model. Importantly, Himba individuals did not sort in line with the “universal” model, even when translated versions of English categories were provided to them (Gendron, Roberson, van der Vyver, & Barrett, 2014). These findings suggest that English language categories may not reflect the categories that are most relevant for Himba participants during facial emotion perception.

Finally, if language helps constitute emotion, then activity in brain regions correlated with language representation, retrieval, and use, should also have increased activity when individuals experience or perceive emotions. Consistent with this hypothesis, meta-analyses of brain activity show considerable overlap in the brain regions with increased activity during language use (e.g., semantics; Binder, Desai, Graves, & Conant, 2009) and those with increased activity during
emotional experiences and perceptions (Kober et al., 2008; Lindquist et al., 2012) (see Figure 2 in Lindquist et al. in press). Shared areas include the dorsomedial prefrontal, ventrolateral prefrontal and temporal cortical areas. Dovetailing with meta-analytic observations, we (Satpute, Shu, Weber, Roy, & Ochsner, 2013) recently manipulated components of emotional experience and observed that areas related to semantics played functionally dissociable roles during emotional experiences. Medial prefrontal regions and temporoparietal areas had increased activity when individuals retrieved mental state categories relevant for making meaning of their body states (i.e. making “I feel” judgments). By contrast, lateral prefrontal regions had increased activity when individuals retrieved specific semantic categories to make meaning of their feelings (e.g. affective labels such as “neutral”, “bad”, “good”). Finally, consistent with the psychological constructionist account that language helps make affect meaningful, we found that limbic/paralimbic regions correlated with the intensity of experienced negative affect.

Together, these findings suggest that words may in fact be constitutive of emotion, suggesting that feelings don’t just get “put into words,” but that words get “put into feelings” too. However, not all models of emotion conceive of language as constitutive of emotion. We turn now to a perspective that views language and emotion as separate systems, but which hypothesizes that language can modulate emotions after they are formed.

**Putting feelings into words: Language and the regulation of emotion**
According research on *emotion regulation* in psychology and neuroscience, language can serve as a means of altering an emotion after it is formed. Broadly, emotion regulation refers to a family of strategies in which people voluntarily increase or decrease the intensity, meaning and/or expression of their emotional experiences (e.g., Beauregard, Levesque, & Bourgouin, 2001; Gross, 1998; Ochsner, Bunge, Gross, & Gabrieli, 2002). Models of emotion regulation often take a dual-systems approach, assuming that the processes involved in emotion regulation are distinct from the processes involved in emotion generation (cf., Gross & Barrett, 2011). The bodily, “bottom-up” processes involved in producing the emotion are thus considered distinct from the cognitive, “top-down” processes involved in regulating the emotion (e.g., Gross, 1998; Ochsner et al., 2002). If language is involved in regulating emotion, then it might be one of the so-called “top-down” mechanisms mediated by cognitive systems of the brain.

One of the most commonly studied cognitive mechanisms of emotion regulation is reappraisal, a method by which people up or down-regulate their affective response to a stimulus by thinking about or reappraising the stimulus or their reaction to it in another way. The extent to which reappraisal involves language is unclear, but it stands to reason that people may be relying on emotion language when they engage in reappraisal. Imagine, for instance, reappraising the feeling experienced when standing atop a skyscraper as “exciting” as opposed to “frightening.” In theory, a person is drawing on knowledge about excitement to do so, and may even be implicitly or explicitly labeling the state anew in his/her own mind.
Many studies reveal that participants can successfully decrease or increase self-reported affective experiences using reappraisal (see Gross, 1998; Ochsner et al., 2004), and are suggestive that language could be involved in this process. For instance, neuroimaging studies find that the act of reappraising a stimulus correlates with activity in the ventrolateral prefrontal cortex and dorsomedial prefrontal cortex, areas that are implicated in the representation of semantic knowledge (Binder et al., 2009) and semantic retrieval (Badre & Wagner, 2007; Satpute, Badre, & Ochsner, 2014; Wagner, Pare-Blagoev, Clark, & Poldrack, 2001). Reappraisal also correlates with decreases in amygdala activity while viewing unpleasant images (for a meta-analysis see Buhle et al., 2013). Since the amygdala is a brain region that responds to the presence of salient or uncertain stimuli and produces autonomic responding (see Cunningham & Brosch, 2012; Whalen, 2007), a reduction in amygdala activity is taken as convergent evidence of successful emotion regulation. These findings imply that language might be involved in reappraisal, but a clear test of the hypothesis that reappraisal depends on language has yet to be conducted.

By contrast, another area of the emotion regulation literature explicitly hypothesizes that language plays a role in emotion regulation. It has been long known that putting feelings into words after the fact can serve as a form of emotion regulation (Pennebaker & Beall, 1986). Pennebaker and colleagues found that writing or talking about one’s emotions can reduce long-term distress associated

---

1 As we pointed out in the first section of this chapter, these same areas are also involved during the experience and perception of emotions (Lindquist et al., in press), and suggest that language may play a constitutive role in the generation of emotional states in the first place.
with traumatic events (Pennebaker, 1997). More recent studies have explored the role of labeling emotions in the moment, or “affect labeling” (cf., Lieberman et al., 2007). Unlike reappraisal, which asks participants to actively reconstruct the meaning of a stimulus, affect labeling involves simply relating a single word with a stimulus (e.g., relating the word “anger” with a picture of a scowling face or feelings in the body). This task does not instruct participants to change the intensity of their emotional state, but nevertheless, labeling one’s state has the unintended or incidental impact of reducing the intensity of emotional experiences (e.g., Lieberman, Inagaki, Tabibnia, & Crockett, 2011).

In one experimental paradigm, participants are asked to match a photograph of a person making an affective facial expression (e.g., a face with wide eyes) with one of two verbal labels presented below it (e.g., “fear” or “happiness”). Lieberman and colleagues propose that accessing words to describe perceptions causes participants to engage in “reflective consciousness,” the kind of consciousness that involves thought and symbolic language use. Engaging in reflective consciousness is thought to simultaneously cause detachment or disruption from “reflexive consciousness,” the type of consciousness that involves focusing on “qualia” from perceptions in the world or feelings in one’s body (cf., Lieberman, 2011). Their hypothesis is that the very act of using language to describe perceptions detaches individuals from the impact of those perceptions. Consistent with this hypothesis, activity in the amygdala was reduced when participants were asked to match an emotional facial expression (e.g., a scowling face) with one of two words (e.g., “anger” v. “disgust”) (Lieberman et al., 2007). The impact of affect labeling has also
been observed behaviorally. Spider phobics took more steps towards a caged spider after using emotion words to describe their situation than when using neutral words (Kircanski, Lieberman, & Craske, 2012).

A current question of interest, particularly for this chapter, is the degree to which affect labeling also has an impact on emotional experience. Most studies of affect labeling present participants with facial expressions (Lieberman, 2011). In these contexts, participants may be labeling the meaning of someone else’s facial expression rather than their own affective experience. Only a handful of studies have examined affect labeling in contexts that are more oriented towards eliciting and measuring emotional experience in the participant. Some of these studies show reduced affective experience when labeling emotional experience (Lieberman et al., 2011), but others do not (Kircanski et al., 2012; McRae, Taitano, & Lane, 2010). Nevertheless, findings from affect labeling are all the more impressive when considering that these studies—relative to reappraisal studies—do not involve the voluntary intention to change emotional states. Rather, the instructions to participants are minimal; they are merely asked to match a single word with a stimulus. Lieberman thus suggests that affect labeling is a form of incidental affect regulation; that when children are told “use your words!,” the mere act of doing so reduces unpleasant affective states.

Affect labeling resembles mindfulness based meditation approaches, which have also been shown to reduce unpleasant feelings. In most forms of mindfulness, practitioners are instructed to label their psychological state using a word (e.g. “anger”) with non-judgmental awareness (Brown, Ryan, & Creswell, 2007): The
mind is trained to observe sensations dispassionately, without exerting motivation to maintain or remove them. Mindfulness has been associated with a variety of health benefits and stress reduction (Grossman, Niemann, Schmidt, & Walach, 2004), but of more relevance here is the resemblance of mindfulness-based techniques to affect labeling. For instance, dispositional mindfulness appears to have an interactive effect with affect labeling: activity in prefrontal cortical regions is greater and activity in the amygdala is less when individuals high in trait mindfulness perform an affect labeling task (Creswell, Way, Eisenberger, & Lieberman, 2007). These findings suggest that individuals high in trait mindfulness may in essence be habitual affect labelers.

In stark contrast to the idea that habitual affect labeling reduces the intensity of one’s emotion is a body of literature from anthropology, linguistics and psychology, which sees language and emotion as fundamentally distinct systems that only interact for the sake of communication. We now turn to this literature, which focuses how emotional experiences and perception get translated (often imperfectly) into the “emotion lexicon.”

**Putting feelings into words: The emotion lexicon**

Research on the “emotion lexicon” systematically describes the terms that speakers of different languages use for emotions (for an excellent example of English emotion categories, see Clore & Ortony, 1988). This approach (with a few notable

---

2 Separate but related lines of research (for a review see Majid, 2012) focus on how affect/emotion is conveyed in classes of words such as interjections (Gendron et al., 2014; Sauter, Eisner, Ekman, & Scott, 2010; Simon-Thomas, Keltner, Sauter, Sinicropi-Yao, & Abramson, 2009), ideophones (Oda,
exceptions (e.g., in linguistics; Pavlenko, 2006; Wierzbicka, 1999) tends to see language as epiphenomenal to emotion. Indeed, the dominant view within this literature is that the capacity to experience and perceive certain emotions is innate and universal; these universal experiences then “sediment” out in language for the sole purpose of communication (cf., Fontaine, Scherer, & Soriano, 2013). This view thus assumes that emotions are “natural kind” categories that consist of a class of universal experiences that are united by a deep causal mechanism and shared surface features (e.g., all instances of anger are similar because they have the same biological mechanism and produce similar observable feelings, physiological changes, and behaviors across instances) (see Barrett, 2006a; Lindquist et al., 2013).

The language sedimentation idea is most likely the dominant perspective in research on the emotion lexicon because it is most consistent with common sense—people essentialize emotion categories (e.g., Lindquist et al., 2013), assuming that emotion words map on to natural kind categories with universal metaphysical essences that make them what they are. People also believe that language is exclusively for communication, despite growing evidence that language feeds back to intrinsically shape mental states ranging from emotions (for reviews see Lindquist & Gendron, 2013; Lindquist et al., in press) to basic visual perception (e.g., Lupyan & Ward, 2013). This said, there are alternate viewpoints, largely in linguistics (Pavlenko, 2006; Wierzbicka, 1999) and in approaches that trace the

2000) and metaphor (Kövecses, 2003; Lakoff, 1987). There is also research assessing the impact of grammar (e.g., syntax), the sound of language (e.g., prosody), and the distinct ways that words are combined in discourse on emotion. This latter topic is an increasing field of inquiry in computing. The growing field of “sentiment analysis” analyzes natural language use (typically in “big data”; Pennebaker, Páez, & Rim, 2013) with the goal of deriving a measure of the subjective state of the individual (be it an attitude, affective state or emotion) (e.g., Kramer, Guillory, & Hancock, 2014).
history of the emotion lexicon (Frevert et al., 2014), that assume, like a psychological constructionist view, that specific emotion categories are not necessarily natural kinds but are social constructions that vary across culture and over time.

Since it is frequently assumed that emotion categories each share a universal “essence,” research on the emotion lexicon likewise assumes that language is a mere “representation” of the emotion categories that already exist, albeit a sometimes imperfect representation of those categories; the translation of emotional experiences and perceptions into words is thus thought to occur in a largely consistent manner across languages. From this perspective, some languages might have an emotion lexicon that is relatively “accurate” and others a relatively “inaccurate” representation of the emotional states that humans are biologically prepared to experience and perceive. Much ethnography on the emotion lexicon has implicitly anchored on this assumption. For instance, Russell (1991) concluded that across 114 ethnographies of emotion surveyed “the ethnographer assumed that the way in which emotion is described in English suited that society and...that native words could be accurately translated into English” (p.433).

In cases where a translational equivalent for an English emotion word does not exist, it is typically assumed that the emotion still exists in nature, but that the society did not develop a need to communicate about that particular state. For instance, (Levy, 1984) referred to states that exist but are not marked with language as “hypocognized,” with the assumption that there are universal patterns of expression/behavior/physiology that mark these states even if they are not
represented in language. In this view, individuals may experience and perceive states (e.g., sadness in Tahitians) that they have no language to communicate. For example, this assumption was tested in speakers of Yucatec Maya, a language that does not have a word for “disgust” (Sauter et al., 2010). The researchers examined whether Yucatec Maya speakers would still be able to differentiate between scowling (“angry”), wrinkle-nosed (“disgusted”), and frowning (“sad”) caricatures of facial expressions (i.e., show categorical perception, or perceptual distortions of a linear continuum of facial actions) despite not having a word for “disgust” in their language. Yucatec Maya participants could reliably differentiate caricatured facial expressions with wrinkled noses from caricatures with frowns or scowls. However, these findings are open to alternate interpretations because participants saw prototypes of each caricature prior to completing the categorization task, during which they received feedback on the accuracy of their judgments. This experience could have allowed them to form perceptual representations for the three different categories even if they did not previously possess separate representations of each of the three categories. At the very least, these findings are inconsistent with other studies in which individuals fail to show categorical perception for emotional faces when they do not know a corresponding emotional word (e.g., Fugate et al., 2010) or cannot access the relevant emotional word due to verbal load (e.g., Roberson & Davidoff, 2000).

Still other research focuses on documenting diversity in emotion language across cultures, focusing on unique categories that do not have a clear translational equivalent across cultures. For example, emotions such as liget and amae are
concepts for which there is no translational equivalent in English. It is clearly
difficult to grasp the meaning of a word that does not have an English translational
equivalent, but researchers have attempted to unpack the meaning of such terms by
describing the contexts in which they are typically used or combining English
language concepts together. For example, *liget* is an emotion that compels members
of the Ilongot tribe to kill others by beheading (Rosaldo, 1980). Based on this
context alone, an obvious English translation of *liget* might be “anger.” Yet (Rosaldo,
1980) also indicates that the “Ilongots see liget in the perspiration of a person hard
at work” and invoke “imagery of focused liget in magical spells before they harvest
rice” (p. 24). As a result it can be argued that no single word, or even simple
combination of emotion terms from the English language lexicon will capture the
meaning of *liget*. Similarly, *amae* is a term from the Japanese language and is the
emotion experienced when you “depend and presume upon another’s love” (Doi,
1973, p. 180)—it occurs in the context of being lovingly cared for. This emotion
word has no exact translation in English, although research (Niiya, Ellsworth, &
Yamaguchi, 2006) shows that English language speakers can understand aspects of
the concept of *amae* when associated with Western situations (e.g., a friend asks you
for help with the computer in the middle of the night). This finding underscores a
caveat about the literature on non-English language emotion categories, more
generally: It is possible that there is a bias towards unearthing and publishing about
the concepts that are most easily imported into English, simply because they are
understandable (even when that understanding comes from combining several
English concepts or anchoring on scenarios typical of Western individuals). For
instance, English speakers may understand *amae* because the concept maps on to some of the situations that they can identify in their own lives (Niiya et al., 2006). That said, English speakers likely have a more narrow understanding of the concept than Japanese speakers, who use the concept to refer to more than just a state that is felt. *Amae* can refer to “an emotion that a person holds toward another person, an interpersonal relationship, a behavior, or even a belief” (Niiya et al., 2006).

Similarly, despite not having a term for *schadenfreude* (pleasure at another’s pain) in English, many English speakers easily understand the word. Indeed, *schadenfreude* has been imported into common English parlance and is now the topic of study by English-speaking researchers (e.g., Cikara & Fiske, 2012). By contrast, there are clearly other concepts that do not possess translational equivalents in English and which English speakers are not readily adopting into daily language, such as *liget* in Ilongot (an exuberence during aggressive acts, described as a force of life; Rosaldo, 1980), *fago* in Ifaluk (sharing features with the English language terms of compassion, love and sadness; C. Lutz, 1988), and *lajja* in India (sharing features with the English language term shame, often occurring in context of publicly aired achievement; Menon & Shweder, 1994).

A lesser-known set of findings focus on cultures in which there is a lack of translation for English language emotion words (for a review, see Russell, 1991). For example, (Russell, 1991) reviewed (primarily ethnology) reports that some cultures lack terms for specific English-language emotion categories. Tahitians appeared to lack a term for *sadness* and *guilt* (Levy, 1973). A term for *guilt* also appeared to be lacking in the Sinhala language of Sri Lanka (Obeyesekere, 1981), the
llongot language of the Philippines (Rosaldo, 1980), the Pintupi language of aboriginal Australians (Morice, 1978), the Samoan language (Gerber, 1975), and in the Ifalukians (C. Lutz, 1980). A term for depression appeared to be lacking in the Yoruba of Nigeria (Leighton et al., 1963), the Fulani in Africa (Riesman, 1986), the Xhosa of South Africa (Cheetham, 1976), the Kaluli of Papua New Guinea (Schieffelin, 1985), in indigenous North American languages (Leff, 1973; Termansen, 1970), in the Malay (Resner, 1970), and Chinese (Chan, 1990). Similarly, the term anxiety is lacking in the Eskimos of North America and the Yoruba of Nigeria (Leff, 1973), the Chinese (Cheng, 1977), and the Machiguenga of Peru (Johnson, 1986). More recently, it was also documented that the Yucatec Maya lack a term for disgust (Sauter, LeGuen, & Haun, 2011). This list is, of course, by no means comprehensive. The strong emphasis on cultural universality in the emotion lexicon literature may mean that there are many more instances in which cultural differences in emotion concepts have not been documented. Strikingly, some researchers who uncovered lack of translational equivalents still interpreted their findings within a universalism framework. For example, Levy (1973) suggested that evidence of crying (a behavior) in the Tahitians was evidence that they still experienced sadness, even if the emotion did not sediment into language.

The lack of simple, single word translational equivalents for emotion categories between English and other languages suggest that English categories may be a limited “anchor” for explorations of the emotion lexicon across cultures. Not only is there variety in the terms for emotion across cultures, but even the overall number of terms in any given lexicon varies widely. According to Levy (1984),
cultures not only “hypocognize” emotions (e.g., the Chewong of Malasia appear to have an emotion lexicon with only seven terms; Howell, 1989), but some cultures “hypercognize” emotions, with thousands of words to mark different states (e.g., over 2,000 terms in the English language). Perhaps even more intriguing is the fact that cultures disagree on what constitutes an “emotion” in the first place—some cultures do not mark “emotions” with a single linguistic category and identify them as a special kind of mental state (e.g. the Samoans; Gerber, 1975; the Gidjingali aborigines of Australia; Hiatt, 1978; the Chewong of Malaysia; Howell, 1989; the Tahitians; Levy, 1973; the Ifalukians of Micronesia; Lutz, 1980; the Bimin-Kuskusmin of Papua New Guinea; Poole, 1985).

At the intersection of extreme universalism and extreme relativism lies a literature that seeks fundamental commonalities amongst emotion lexicons. This research began with Osgood and colleagues, who assessed core commonalities that define emotion lexicons across cultures during the 1950s and 60s. Osgood (1975) employed a semantic differential approach in which participants rated the meaning of emotion terms on a number of different bipolar adjective scales. For example, scales might range from “good” to “bad,” or “strong” to “weak,” with neutral as the mid-point. Using data reduction techniques, Osgood revealed that three dimensions—evaluation, potency, and activity—contributed to the connotation of emotion words across 20 different cultural contexts (Osgood, 1975). Since this seminal work, the same or similar dimensions (e.g., valence and arousal; Russell, 1983) have been documented across many different studies, using a variety of methods (for a review, see Russell, 1991).
More recently, Fontaine and colleagues again assessed the dimensional space that best accounts for the meaning of emotion terms across multiple cultural contexts (Fontaine, Scherer, Roesch, & Ellsworth, 2007). Their approach was theoretically anchored in a “componential” framework for emotions. The authors thus asked participants to rate emotion words on a set of priori scales assessing action tendencies, subjective experience, and regulatory aspects of emotion (i.e., the GRID instrument; Scherer, 2005). Dimension reduction of data collected across Belgium, the United Kingdom, and Switzerland revealed four dimensions that characterized the similarities between emotion terms: evaluation, potency, arousal, and unpredictability (Fontaine et al., 2007). Although these dimensions are quite similar to those derived in prior research (particularly the first three dimensions, which are largely consistent with Osgood’s results), the fourth dimension of unpredictability was unique to Fontaine et al.’s findings (Fontaine et al., 2007; Fontaine et al., 2013). One explanation of the discrepancy between Fontaine et al.’s findings and the large literature that observes only two or three dimensions is methodological. Fontaine et al. used unipolar rating scales, whereas previous dimension reduction research used bipolar rating scales (for review see Russell, 1991). The nature of the scale can impact the correlations between items in factor analysis, with unipolar scales decreasing correlations between items and inflating the number of independent factors observed (for a discussion see Russell & Carroll, 1999). There is justification for using bipolar scales because participants often implicitly impose a bipolar opposite on a unipolar scale, even if the end point is specified as “neutral” (e.g., they assume a scale ranges from “unpleasant” to
“pleasant,” even when the scale specifies “neutral” to “pleasant”) and this causes measurement error (for a discussion see Russell & Carroll, 1999). Second, prior research did not invoke an a priori set of “components” for emotions. As a result, the structure underlying emotion terms may differ when perceivers anchor on their internal representations of emotion in a relatively unstructured manner versus when cued about specific content as in Fontaine et al. (2007). For example, items related to the “regulation” of emotion appear to load highly on to the “unpredictability” dimension in Fontaine et al’s analysis, but it’s not clear that individuals standardly think of the regulatory implications of an emotion when they are not prompted to emphasize that meaning of an emotion term.

Wierzbicka’s Natural Semantic Meta-language (NSML) offers another approach to understanding the commonalities between emotion terms. Wierzbicka’s linguistic research indicates that there is minimal universality in concepts across cultures and she argues that more fundamental concepts (e.g., “good”, “bad”, “do”, “happen”, “know”, etc.), should form the foundation for cross-cultural comparisons of emotion (Wierzbicka, 1999, 2009). The NSML approach is a useful tool for revealing whether words that are assumed to be translational equivalents in different languages (e.g., Russian “smertnaja muca” and English “sorrow,” Wierzbicka, 2009) really have the same underlying meaning. This approach has not been adopted into mainstream psychology (Wierzbicka, 2009), however.

Of course, cultural diversity is not the only factor that produces differences in emotion lexicons. A growing literature also explores the development of emotion
words across the lifespan (see Widen, 2013). For instance, children follow a slow
and fairly predictable trajectory of linguistic emotion category acquisition (e.g.,
Widen, 2013). This development is striking because it suggests that some emotion
terms, and the corresponding perceptual representations of that emotion category,
emerge much later than others. For instance, two-year-old children reliably use the
words “happy” and “angry” to describe faces and can correspondingly distinguish
positive from negative faces. However, children do not become able to meaningfully
use the words “fear” until later in childhood (around 4 years of age) and do not
reliably differentiate fearful faces from other negative faces until this point (see
Widen & Russell, 2008). As we note in the next section, such findings are ultimately
consistent with the psychological constructionist view, in which language helps to
constitute emotion by driving category acquisition and online categorization of
one’s experiences or others’ facial expressions.

**Summary**

Throughout this chapter, we reviewed three traditions that each explore the
relationship between language and emotion. The psychological constructionist
approach assumes that concept knowledge supported by words constructs
emotions in the first place, when affect is made meaningful as an instance of an
emotion category (that is relevant to the speaker of a given language). By contrast,
the emotion regulation research assumes that words can feedback to modulate
emotions, helping to regulate emotions after the fact. Finally, the work on the
emotion lexicon typically assumes feelings merely get put into words after the fact.
The fact that so many approaches are concerned with the relationship between
language and emotion underscores the importance of this topic area in contemporary psychology, neuroscience, linguistics and anthropology. However, the three approaches we discussed herein appear to be mechanistically inconsistent with each other, at least on the surface. We argue that despite their apparent inconsistencies, a psychological constructionist approach can in fact unite findings from across these seemingly diverse domains.

For instance, if language helps to constitute emotional feelings in the first place by shaping how people make meaning of affective states, then it follows that prompting people to make meaning of their states with linguistic categories (as in affect labeling) or to re-construe the meaning of a feeling with a different linguistic category (as in reappraisal) will contribute to the regulation of emotions. Of course, there is debate about whether the processes involved in emotion regulation are the same or different than those involved in emotion generation (Gross & Barrett, 2011), but early evidence is suggestive that the neural mechanisms involved in emotion experience and regulation are similar (cf., Ochsner, Silvers, & Buhle, 2012). Furthermore, there are alternative interpretations of the affect labeling literature that are consistent with the psychological constructionist view. For instance, it is possible that putting feelings into words actually forces individuals to make meaning of their otherwise ambiguous feelings of affect (pleasantness/unpleasantness or activation/deactivation) towards stimuli (cf., Lindquist et al., in press). For instance, the finding that explicitly labeling facial expressions with emotion words decreases activity in the amygdala (Lieberman et al., 2007) could be evidence that words help reduce the uncertainty of affective
stimuli. The amygdala, although broadly involved in emotion, is thought to be specifically involved in signaling the brain to process uncertain stimuli further (Whalen, 2007). As such, an alternate, but not incompatible, interpretation of the affect labeling findings is that language helps to regulate emotion by reducing uncertainty about the meaning of sensations in the body (as in emotion experience) or the world (as in emotion perception). Once a person constructs an emotional experience or perception and knows what his or her sensations mean, he or she can successfully regulate them. Indeed, greater specificity about the meaning of one’s emotions in daily life (called “emotional granularity”) is associated with greater emotion regulation success across instances (e.g., Barrett, Gross, Conner, & Benvenuto, 2001), perhaps because knowing the meaning of one’s affective state makes it easier to regulate that state (Kashdan, Barrett, & McKnight, in press; Lindquist & Barrett, 2008b). Indeed, emotional intelligence interventions in children first teach children about different emotion concepts (e.g., anger, fear, disgust, etc.) and how to differentiate between instances of those concepts. Once knowledge about the concepts has been learned, children are taught how to regulate their feelings (Rivers, Brackett, Reyes, Elbertson, & Salovey, 2013). The idea that the ultimate mechanism of affect labeling is the categorization of an affective state as an instance of emotion is more consistent with psychological constructionist accounts of emotion, than the assumption that language only feeds back to shape emotions after the fact. Of course, more research is needed to better understand whether the role of language differs in emotion regulation vs. emotion construction.
Just as the emotion regulation literature is ultimately consistent with a psychological constructionist view, the emotion lexicon literature can be interpreted as consistent with the idea that language helps construct emotions in the first place. For instance, if words in part constitute emotions, then this might describe why different cultures possess different emotion categories, why there are imperfect translations between different languages about the meaning of emotion words, and why children do not have fully formed emotion concepts until they know the relevant emotion category words. We argue that instead of revealing imperfect linguistic concepts for natural kind categories that already exist in the world, research on the emotion lexicon instead reveals the conceptual content that has the power to shape emotions in the speakers of that language. Consistent with the psychological constructionist view, the emotion categories represented in language might thus be better considered “cognitive types” (Clore & Ortony, 2008, 2013), “nominal kinds” (Barrett, 2012; Lindquist et al., 2013) or situation-specific concepts (Wilson-Mendenhall et al., 2013; Wilson-Mendenhall et al., 2011) that have the power to shape how speakers of a language make meaning of affective experiences or perceptions in a given moment. This interpretation would explain why speakers of different languages literally “see” different emotions on faces (Gendron et al., 2014; Jack et al., 2012) or describe instances of the same emotion category differently (Wierzbicka, 2009). Until only very recently (Gendron et al. 2014b), researchers assumed that translations of English emotion categories (if they exist) are the same categories that are used by all cultures to describe day-to-day emotional experiences and perceptions. Yet the findings of Gendron et al. (2014a, b)
suggests that when studying emotional experiences and perceptions across cultures, researchers should use more open-ended methods to discover (rather than stipulate) what categories members of a culture spontaneously uses to make meaning of affective experiences and perceptions in their own daily lives.

In total, the research reviewed in this chapter suggests that conceiving of words as constitutive of emotion charts a new path forward for the science of emotion, helping to unite seemingly disparate traditions of study and suggesting new implications. Taking into consideration that words get put into feelings, alongside the more commonsense notion that feelings get put into words, might just change what researchers discover about the role of language and emotion.
References


Kassam, K. S., & Mendes, W. B. (2013). The effects of measuring emotion: physiological reactions to emotional situations depend on whether someone is asking. PLOS One, 8(7), e64959.


Satpute, A. B., Badre, D., & Ochsner, K. N. (2014). Distinct regions of prefrontal cortex are associated with the controlled retrieval and selection of social information. *Cerebral Cortex, 24* 1269-1277.

