Sensory language across lexical categories

Winter, Bodo; Strik Lievers, Francesca

DOI:
10.1016/j.lingua.2017.11.002

Sensory language across lexical categories

Francesca Strik Lievers[1] & Bodo Winter[2]

[1] University of Genova,
Dipartimento di Lingue e Culture Moderne

[2] University Birmingham,
Department of English Language and Applied Linguistics

Abstract

Being able to talk about what humans perceive with their senses is one of the fundamental capacities of language. But how do languages encode perceptual information? In this paper, we analyze how experiences from different senses (sight, sound, touch, taste, and smell) are encoded differentially across lexical categories (nouns, verbs, adjectives) in the English language. Three independently collected lists of perception-related words show that sound concepts are more prone to being expressed as verbs. Data from an independent rating study furthermore shows that nouns rated to strongly relate to motion are also rated to strongly relate to sound, more so than is the case for color-related nouns. We argue that the association of verbs with sound is due to sound concepts being inherently more dynamic, motion-related and event-based, in contrast to other sensory perceptions which are phenomenologically less strongly associated with motion and change. Overall, our results are the first to show differential encoding of perception-related concepts across different types of lexical categories. Our analyses of lexical patterns provide empirical evidence for the interconnection of semantics and grammar.

Keywords: audition; sound; perception; parts-of-speech; nouns; verbs

1. Introduction

Humans perceive the world through their senses and then share their perceptions with others, chiefly through language. Talking about sensory perceptions, such as whether a curry tastes too spicy or a fish smells rotten, forms a frequent focus of communicative activity. How people talk about sensory perceptions has been studied for a long time in linguistics (e.g., Ullmann, 1959 [1957]; Williams, 1976; Viberg, 1983; San Roque et al., 2015), anthropology (e.g., Classen, 1993; Howes, 2003), and the cognitive sciences (e.g., Miller & Johnson-Laird, 1976; Majid & Burenhult, 2014; Olofsson & Gottfried, 2015; Majid, 2015). A core question in this area of research is how perceptual meaning is encoded in the lexicon of human languages (Levinson & Majid, 2014). What words are available to speakers from various languages to talk about what they see, hear, feel, taste and smell?

Research that looks at how perceptual experience is expressed in language is theoretically and methodologically heterogeneous, but it is possible to observe that it has been conducted from two main perspectives. The first perspective takes a specific lexical category as the starting point. The second one takes sensory modalities as a starting point. In the lexical category perspective, studies focus on a given lexical category—often either verbs or adjectives—and observe the interplay between sensory modalities within that lexical category. For instance, typological research on verbs of perception has shown that most languages have more verbs for vision and/or for hearing than for the other modalities; moreover, verbs of vision and/or hearing may undergo a semantic extension toward the other modalities, while the reverse happens less frequently (Viberg, 1983, 2001; Evans & Wilkins, 2000; Maslova, 2004; Vanhove, 2008). As for perceptual adjectives, research has focused on semantic extensions, particularly highlighting the fact that certain sensory modalities are more likely to be associated with each other in adjective-noun pairs than others. For instance, touch-related adjectives are often used to modify hearing-related nouns (e.g., *rough*: *rough voice*, *rough sound*), but it rarely happens that hearing-related adjectives are used to describe tactile perceptions (Ullmann, 1959 [1957]; Williams, 1976; Shen, 1997; Ronga et al., 2012; Strik Lievers, 2015; Winter, 2016b; Ronga, 2016). The empirical studies of perception verbs and “synesthetic” adjectives mentioned here are examples of studies that primarily focus on a particular lexical category or combinations of particular lexical categories (as in adjective-noun pairs).

The sensory modality perspective focuses on one (or more) sensory modality, investigating the characteristic way the sensory modality/modalities are encoded in the vocabularies of human languages. Many such studies concern, for instance, the fact that different sensory modalities seem to be linguistically encodable to different degrees (Levinson & Majid, 2014; Majid & Burenhult, 2014; Olofsson & Gottfried, 2015), both within a single language and across languages/cultures. In English and other Indo-European languages the expression of visual perceptions can rely on a particularly rich vocabulary compared to that available for the other senses (Buck, 1949). On the contrary, smell has a very small number of dedicated lexemes: it is presumed to be the most “ineffable” sense (cf. Levinson & Majid, 2014). While visual language appears to be dominant not only in Indo-European languages, but also in all the other languages that have been analyzed so far in the literature, the ranking of the other senses seems to be more variable cross-linguistically and cross-culturally (Majid & Burenhult, 2014; San Roque et al., 2015; O’Meara & Majid, 2016).

Both the lexical category and the sensory modality perspectives have been widely studied. In this paper we explore a third perspective, which has so far received less attention, and which can be seen as a combination of the other two. We address the following research question: are there differences among lexical categories in the expression of concepts pertaining to the various senses? To put it another way: how many words of a given lexical category are there for a given sensory modality, such as sight, sound, touch, taste, or smell? That is, do the different sensory modalities differ in which *kinds of words* are preferentially used to describe them? For instance, Strik Lievers (2015: 86-88) observed that, in her English and Italian datasets of sensory lexemes, adjectives are numerous for touch and few for hearing, while nouns are abundant for hearing and scarce for touch. Knowing about this distributional fact of the sensory lexicon may explain, in part, why adjective-to-noun mappings are more commonly touch-to-sound rather than sound-to-touch, as has been previously described by researchers working on synesthetic metaphors. In other words, the directionality observed in metaphorical mappings between sensory words may be related to an already existing asymmetry in the lexicon (for a similar idea, see already Ullmann, 1959 [1957]: 283; see also Winter, 2016b: Ch. 8).

This paper investigates the distribution of sensory lexemes across lexical categories systematically by comparing different datasets of sensory lexemes that have been built for English in previously published studies. A quantitative analysis shows that the senses indeed differ with respect to how many verbs, adjectives and nouns they have. We argue that this asymmetric distribution can be related to the different properties of prototypical representatives of the various lexical categories on the one hand, and to phenomenological and perceptual differences between the senses on the other. That is, we identify differences between the senses that can be related to semantic differences between lexical categories.

Reasoning about the motivations that may explain the distribution of sensory lexemes across lexical categories will hopefully shed new light on the connections between the characteristics of actual human perception and the linguistic means used to express it. In addition, the results of this work contribute to explaining the tendencies that have been observed for synesthetic transfers (Strik Lievers, 2015) and other meaning extensions, such as the observation that perception verbs frequently extend their meaning to encompass the other senses (Viberg, 1983; Evans & Wilkins, 2000; see also Sweetser, 1990). To interpret these proclaimed results about how the senses differ in language one needs to have an understanding about what the baseline frequency of sensory words is with respect to particular lexical categories. Finally, in light of the fact that semantic criteria differentiating adjectives, nouns and verbs are often not deemed as important compared to distributional tests (see discussion in Baker & Croft, 2017), our empirical results provide an important foray into characterizing the semantics of English lexical categories in a quantitative fashion. In addition to our descriptive and theoretical contributions to the study of language and perception, and to the study of lexical categories more generally, we also make a methodological contribution by showing how existing databases—in particular norm datasets with ratings collected by humans—can be used for linguistic theorizing in this domain. Many claims that have been made in the past without quantitative substantiation can now be addressed using already published databases (see also Winter, 2016a, 2016b).

2. Background on lexical categories and the senses

2.1. Lexical categories

One of the core properties of language that has received much discussion in formal linguistics, functional and cognitive linguistics, as well as in typology, is that words in the lexicon are grouped in what are often called “parts-of-speech” or, with a more or less overlapping meaning across the literature, “syntactic categories”, “word classes”, “grammatical classes” and “lexical categories”, the latter being the label used here. As outlined for instance in Givón (2001: Ch. 2), there are multiple criteria that can be used to define lexical categories: morphological criteria (which types of affixes attach to the lexical root), syntactic or distributional criteria (which slots in phrases are occupied by the word; see Berg, 2000, for discussion), and semantic criteria (the types of meanings words encode) (see Croft, 1991 and Rauh, 2010 for additional criteria and a review of theoretical approaches, and Baker & Croft, 2017 and Himmelmann, 2017 for further references). In addition, there is a whole set of phonological criteria (Sherman, 1975; Liberman & Prince, 1977; Cassidy & Kelly, 1991; Kelly, 1988, 1992; Kelly & Bock, 1988; Taylor, 2002; Monaghan et al., 2005; Hollman, 2013), which distinguish nouns and verbs in a probabilistic fashion. Here, we focus on the three major categories: nouns, adjectives and verbs, arguably the three most important major content classes (Baker, 2003).

In English, there are several morphological criteria that help to distinguish the different word classes from each other. For example, English nouns have plural and genitive affixes (*chair*s, *chair*'s), while words of other lexical categories do not; but this only holds for some nouns (not for mass nouns as *water*, for instance). In fact, morphology cannot be used as a unique defining criterion, and needs to be combined with syntactic or distributional criteria. Nouns, for instance, may immediately follow a determiner and/or an adjective (as in *a white chair*, *blue water*), while words of other lexical categories cannot (Aarts & Haegeman, 2006: 118). In connection to the reduced morphology that characterises the English language, it is moreover not uncommon that the same word form can be said to belong to different lexical categories depending on the context it is used in. Compare "I ordered a *drink* (noun)" with "I *drink* (verb) too much coffee" (Ježek, 2016: 101). Of course, the two *drink* differ not only in their distribution, but also in their semantics, an issue to which we return below.

Another formal set of criteria that differentiates nouns from verbs alongside morphosyntactic criteria are phonological criteria. Numerous studies have shown that nouns and verbs do, in fact, *sound* differently. For example, English verbs have a statistical tendency to end in final voiced stops (Taylor, 2002; Monaghan et al., 2005; Hollman, 2013). Hollman (2013) found that when participants generate nonce verbs and nouns, the nonce verbs had more final obstruents than the nonce nouns. Lexical stress is an important cue to the noun/verb distinction in English (Liberman & Prince, 1977; Kelly, 1988, 1992; Kelly & Bock, 1988), with disyllabic nouns having initial stress, as opposed to disyllabic verbs, on which the second syllable tends to be stressed (see also Sherman, 1975). Nouns also have on average more syllables than verbs (Cassidy & Kelly, 1991). These studies show that phonological patterns help to differentiate nouns and verbs on top of morphosyntactic patterns. The phonological patterns are, crucially, probabilistic: a verb, for example, may well not have a final voiced stop and still be a verb—but verbs are statistically more likely than nouns to exhibit this pattern (see Kelly, 1992; Monaghan et al., 2005, 2007; Farmer et al., 2006). The degree to which a noun or a verb fits the phonological patterns observed for its lexical category is best seen as a prototype category, with some more prototypical nouns and verbs and some less prototypical nouns and verbs. In the case of sound patterns, this prototypicality is determined by how many phonological features of the lexical category apply to a given word (Monaghan et al., 2005).

The same way that we cannot reason about the sound patterns of nouns and verbs in terms of crisp categorical distinctions, the same way we cannot reason about the semantics of nouns, verbs and other lexical categories (Lyons, 1977: Ch. 11) in an absolute fashion. It is widely agreed that “one cannot simply look at the part of speech of a word in order to know its semantic type” (Murphy, 2010: 144). That is, semantic criteria for what makes a verb a verb and a noun a noun are not sufficient to classify all words according to lexical category (but see discussion in Gärdenfors, 2014: Ch. 6-11). For example, the noun *running* describes an action, whereas the verb *to know* describes a state, showing that we cannot use the semantic distinction between actions and states as a hard rule to distinguish nouns and verbs. However, even though lexical categories are reliably defined morphologically and distributionally (and to some extent also phonologically), different lexical categories also have semantic prototypes, which correspond to the ontological categories THING, ACTION and PROPERTY (Murphy, 2010: Ch. 7; see also Frawley, 1992: Ch. 3-4).

Murphy (2010: 144) states that: “[t]ypical nouns describe INDIVIDUAL PHYSICAL OBJECTS, typical verbs describe PHYSICAL ACTIONS, and typical adjectives designate PROPERTIES.” As expected given the probabilistic nature of semantic patterns in lexical categories, examples that do not conform to these prototypes abound, such as nouns that do not denote things (e.g., *oblivion*, *unconventionality*, *examination*), verbs that do not denote actions (e.g., *to know*, *to be*) or adjectives that do not denote properties (e.g., *corporate* in *corporate tax*). Murphy (2010: 140) goes on to make the following statements with respect to the prototypes of nouns, verbs and adjectives:

- “The central members of the NOUN category designate types of PHYSICAL OBJECTS, and PHYSICAL OBJECTS are most typically designated by nouns.”
- “The central members of the VERB category designate types of ACTIONS, and ACTIONS are most typically designated by verbs.”
- “The central members of the ADJECTIVE category designate types of PROPERTIES, and PROPERTIES are most typically designated by adjectives.”

Givón (2001), and following him Murphy (2010), lists several criteria for each of the major lexical categories (see also Frawley, 1992). Similar to Monaghan et al. (2005)’s prototypicality measure for phonological features of nouns and verbs, a word can also be a more or less prototypical member of a lexical category depending on how many criteria it satisfies. Prototype categories are inherently probabilistic, there are no hard cut-off criteria but fuzzy boundaries (this differs starkly from the treatment of lexical categories in formal grammars).

Even only with respect to semantic criteria, Givón (2001: Ch. 2) lists multiple dimensions as well, including temporal stability, complexity, compactness, countability, agentiveness and spatial diffuseness. Of these, Givón (2001) considers temporal stability the most important one (see p. 50). Compared to adjectives and in particular to verbs, "[t]he prototype of the class *noun* occupies the most time-stable end of the scale" (Givón, 2001: 51). As Murphy (2010: 141) states, "[y]ou can be fairly sure that something that is called *a table* will still be a table (and not a goose or a song) from one moment to another". Time-stability is not an all-or-nothing property — there is a continuum from very stable to very unstable (Murphy, 2010: 141; Givón, 1984: 55), so there are gradations of time-stability within the noun category, and for the verb and adjective categories as well. Givón (2001: 51) gives the following examples from the nominal domain:

> "If it is a *chair* now, it is still likely to be a *chair* in five minutes, an hour, or a day — in size, shape, color, texture, consistency or usage. Of course, a fine internal gradation still exists, so that a *child* may change faster than a *tree*, and that faster than a *house*, and that faster than a *rock*, etc."

In contrast to most nouns, prototypical verbs "are not so time-stable — they tend to represent that which is temporary and changing", for example, we do not expect a 'singing' event to last forever (ibid. 141). The idea of nouns and verbs differing with respect to the dimension of time is already found in Aristotle's distinction between *onoma* and *rhēma*: "By a noun we mean a sound significant by convention, which has no reference to time" (*De Int.* 16a, 19-21); and "A verb is that which, in addition to its proper meaning, carries with it the notion of time" (*De Int.* 16b, 6-9) (from Blevins, 2012: 377). In Langacker's term, verbs realize the conceptual schema called "process", which describes "a complex relationship that develops through conceived time" (Langacker, 2008: 112).

Adjectives rank in between verbs and nouns with respect to time-stability. They express concepts that are less stable in terms of their temporal profile than concepts expressed by nouns because they can refer to properties of objects that can change, as well as to properties that may not change. And adjectives express more time-stable concepts than verbs, since they frequently refer to concepts that do not involve rapid changes—e.g., a green apple changes color only slowly, and some objects, such as rocks, do not change color at all, unless painted. This renders adjectives such as *blue* or *yellow* relatively more time-stable than prototypical verbs such as *run* and *throw*. However, again, there are gradations within the adjective category. For example, color adjectives such as *blue* or *yellow* describe relatively more time-stable properties than adjectives describing emotional states such as *happy* or *sad*, which describe properties that can change very rapidly. The issue that adjectives can be both stable or instable in terms of their temporal profiles is also explored in the formal literature, which has discussed at great length the difference between individual-level predicative adjectives (associated with temporal persistence) and stage-level predicative adjectives (associated with bounded states) (see e.g., Marín, 2010). Whereas in the formal literature this distinction is seen as hard-cut, time-stability is conceived of as a graded notion in cognitive semantics approaches, with variation within lexical categories.

Another way to think about these semantic prototypes for nouns, verbs and adjectives is to think from the perspective of lexical differentiation, in line with what has been said about the codability of certain semantic domains (compare Levinson & Majid, 2014): we may think of the semantic domain of actions (and more generally events) as being more differentiated in the verbal category than in the noun category. This view is expressed by Frawley (1992: 68) who says about nouns that they “may not always be persons, places, or things, but persons, places, and things almost always turn out to be nouns”. He similarly says that “[n]ot all verbs are actions, but when actions are expressed, they overwhelmingly tend to surface as verbs” (ibid. 141). According to this proposal (which so far has not been tested quantitatively), for action-related and event-related concepts there should be more verbs than nouns and adjectives. The semantic domain of properties on the other hand should be more differentiated within the adjective category, with (relatively) more adjectives denoting properties, compared to nouns and verbs. Similarly, the semantic domain of objects should be most differentiated within the nominal domain.

2.2. Lexical categories and the senses

Based on these ideas, we can make explicit predictions with respect to the lexical differentiation of sensory words. In particular, the relatively more “dynamic” sensory modalities, i.e. those that are more event-oriented and time-varying, should be more differentiated within the verbal domain because verbs, according to the positions outlined above (Frawley, 1992; Givón, 2001; Langacker, 2008; Murphy, 2010; Gärdenfors, 2014), should load heavily onto those semantic domains that involve transient phenomena such as actions, events and movements.

We propose that the sensory modality of sound (audition) is a prime candidate for a perceptual quality that is inherently dynamic. Auditory experience involves a strong component of “spatio-temporal dynamism” (O’Shaughnessy 2009: 117; see also O’Callaghan, 2007, 2009, 2014 and references therein), arguably stronger than that involved in visual, gustatory, olfactory and tactile experience. There are two aspects of the idea that sound is comparatively more dynamic. First, motion is necessarily involved in sound production, which in many cases results from deliberate actions, and sound production itself is an event that unfolds over time. Second, the sounds we frequently hear are generally transient and if they are not, they involve internal variation such as changes in frequency. Even a “static” frequency of, say, 440Hz, is something that can only be perceived by hearing multiple pulses and integrating them over time. That is, time is an inherent feature of sound in both production and perception. As stated by Matthen (2010: 79-80), “audition presents its objects as *temporally composed*”.

Regarding the action and movement components of sound, consider that a rock by itself does not make noise. But throw a rock and it makes a clonking sound; scratch along its surface and you get a screeching sound; crack the rock and it makes a cracking sound. That is, any action performed on the rock creates sound. The rock itself is static and soundless, but once subject to movement, sound waves are created (“[t]he generation of sound always originates in mechanical vibration”, Hartmann, 1995: 1). We may associate movement of the rock with sound both through our own action (such as throwing the rock), or through external or inanimate action, such as when a rock falls down a cliff because of wind. This is nicely expressed in the following quote from Aristotle’s *De Anima* (book II Ch. 8, 419^{b}9, transl. by D. W. Hamlyn 1968):

> “Actual sound is always of something in relation to something and in something; for it is a blow which produces it. For this reason it is impossible for there to be sound when there is only one thing; for the striker and the thing struck are different. Hence the thing which makes the sound does so in relation to something; and a blow cannot occur without movement.”

O’Callaghan (2009: 28) also states that “sounds are particular events of a certain kind. They are events in which a moving object disturbs a surrounding medium and sets it moving”. Philosophers have extensively discussed the perceptual nature of sound (O’Callaghan, 2014). According to the event-based view of auditory objects of perception, sounds are events that occupy time (in sequence), in contrast to objects (and their properties) which may exist wholly at a particular moment in time (O'Callaghan, 2007, 2008, 2009; Matthen, 2010). O'Callaghan (2008: 804) describes sounds as being different from "ordinary tables and chairs – you cannot grasp or trace a sound – and sounds are not heard to be properties or qualities of tables and chairs, since sounds do not seem bound to ordinary objects in the way that their colors, shapes, and textures do". Even philosophers who contend that sounds are stable dispositions of objects (to vibrate in a certain way in response to the right kind of mechanical stimulation) acknowledge a crucial component of movement in order to make the sounds that an object is disposed to create audible: "We only hear sounds when objects are stimulated to vibrate and thus produce pressure waves in the ambient air. Without stimulation, or without air, you can't hear objects" (Kulvicki, 2015: 207).

As an example where the connection between sound production and action is felt particularly strongly in our everyday phenomenology, consider speech. Vocal production involves movement, such as movement of the diaphragm and the lungs to generate air flow; movement of the vocal folds to generate voiced sounds and pitch; as well as movements of the tongue and the jaw, often accompanied by external bodily movements such as head movements to index prosody or gestures. Another example of the inherent sound/motion connection is gait. We are used to our own movements generating sounds while walking.

Of course, the overlap between movement in our everyday environment and sound is not perfect, at least not when seen from the perspective of our auditory phenomenology. Although sound *necessarily* involves motion in its production, we do not always witness the motion as such. As stated by Pasnau (2000: 34), "one can perceive motion without perceiving sound; (...) one can perceive sound without perceiving motion". Examples of this include seeing movement at a distance, too far away for any sounds to be audible (as often happens with airplanes in the sky), or seeing small insects fly around whose movements are simply too quiet to be audible. Listening to music through ear plugs with one's eyes closed is another example of movement and sound being decoupled in our environments. In the case of ear plugs, there still is movement involved (the vibrations of the sound-emitting device), but we are not phenomenologically aware of these movements. While all movement generates sound and all sound is generated by movement, not all sound and sound-generating movement is available to our phenomenology. That said, since sounds necessarily involve movement in their production and since any physical action necessarily produces some sound, the correlation between sound and movement is particularly strong. This correlation is much weaker for all of the other senses. A flower, for example, does not have to move in order to make the percepts ‘red’ or ‘fragrant’ available to the perceiver; these percepts are accessible without movement.

The other side to the notion that sound is inherently more dynamic is its transient nature (cf. O’Callaghan & Nudds 2009 and Nudds 2015 for recent philosophical perspectives on the essential temporality of sound). As said by O’Callaghan (2007: 22), a sound “has a beginning, a middle, and an end”. As a result of this, we expect sounds to not last forever, such as the events described by words like *squealing*, *beeping, barking* and *clonking*. Even when we use seemingly more time-stable auditory adjectives, such as *loud* and *quiet*, these terms are either bound to a transient sound (*a loud beep*) or can possibly change state (*a quiet classroom*). This is different from, for example, color terms, which describe relatively more stable properties of entities[1]. A *gray rock* will generally change color less quickly than a *quiet classroom* will change quietness. Sound adjectives have a fundamentally different flavor, compared to adjectives of the other senses. For sensory perception one can see, feel, taste or smell, adjectives actually denote properties. In the case of sensory perceptions we hear, adjectives denote events, and these events are often connected with actions (such as *squealing*, which suggests an animate producer). Even if sound is not dynamic in terms of transience of the entire sound (on/off), then it is still characterized by internal variation and by a recognizable internal structure that is temporally defined. Even relatively stable sounds created by inanimate things, such as the sound of the ocean, involve internal temporal patterns—in fact, the notion of perceiving a sound of a certain frequency involves continuous variation in sound waves. Without frequency pulses or some form of rhythmic cycle, we would not be able to hear a sound of a certain frequency.

Thus, sound is two-fold dynamic: it involves actions and movements in its production, and it is dynamic in perception as well, by virtue of being transient and characterized by internal change. This dynamicity is largely accessible to our phenomenology, that is, we are or can become aware of it. Based on this and based on Givón’s notion of time-stability, we can form the prediction that sound concepts should be *relatively* more differentiated in the verbal domain. That is, compared to the other senses, sounds should be more verby in their linguistic patterning.

We can formulate an additional prediction with respect to the modality of touch. It too is very dynamic (see Popova, 2005), but in a different way from sound. Here, we have to differentiate between the perception of *surface texture* as opposed to such

[1] As commented by one reviewer, adjectives like *loud* and *quiet* can also be distinguished from adjectives like *gray* in terms scalarity (locationality vs. configurationality, cf. Popova 2005). In the present discussion we are however concerned with the temporal dimension.

touch-related properties as temperature. Whereas it is possible to perceive heat or coldness by touching an object statically (such as holding one's finger on a stove), the perception of such surface properties as roughness, smoothness, hardness, or softness, is very limited without moving one's hand along a surface. Carlson (2010: 248) mentions that "[u]nless the skin is moving, tactile sensation provides little information about the nature of objects we touch", and Bartley (1953: 401) says that "tactile exploration is a piecemeal affair". This also relates to the distinction between active and passive touch (Katz, 1989 [1925]): "In one case the impression on the skin is brought about by the perceiver himself and in the other case by some outside agency […] Active touch is an exploratory rather than a merely receptive sense. When a person touches something with his fingers he produces the stimulation as it were. More exactly, variations in skin stimulation are caused by variation in his motor activity" (Gibson, 1962: 477). As Popova (2005: 401) points out, "because of the hand's function in active touch, the tactile sense is a unique modality in which stimulation is *obtained* rather than imposed by the stimulus".

The same goes for the perception of shapes via touch. Imagine perceiving the shape of a walking stick. Vision makes the shape percept available to one's consciousness at an instant (see Stokes & Biggs, 2015). In contrast, perceiving the same shape via touch (without sight) involves moving one's hand along the stick and only after having haptically explored the stick for a long time does the full shape become apparent. However, while the perception of surface texture and shape via touch in the absence of vision necessarily involves movement and action, touch may also be slightly less dynamic than sound, or dynamic in a different way. Above we argued that the dynamicity of sound is a two-fold idea, one aspect being the movement dimension involved in sound production, the other one being transience. The action-component also ascribes to (active) touch, which involves haptic exploration and hence movement. However, an important difference here is that sound is dynamic the way sounds are produced and perceived, whereas touch is dynamic only with respect to the way humans perceive. The surface properties and shapes themselves are not the outcome of dynamic events the same way they are for sound. In fact, surface properties are generally more stable properties of objects. For example, a rough rock generally stays a rough rock the same way that its color stays the same. Thus, touch may not be as dynamic as sound.

Of course, it is trivially true that all sensory perception has an element of dynamicity. All of cognition and perception takes time (Spivey, 2007), even the perception of color, tastes or smells. However, what we talk about here in terms of dynamicity is not so much just the phenomenological characteristic of an individual sensory impression slowly unfolding in time, but also the immediate association of particular perceptions with actions and movements, as well as with rapid change and sequential temporal structure. Moreover, it is also trivially true that all perceptions necessarily involve some form of movement, such as the decomposition of chemicals in the case of taste and the movement of light photons in the case of vision. However, the issue is whether the motion is actually *perceptible*. Huumo (2010: 59) (see also Talmy, 2000: 112) notes that the dynamic component of sound (and to some extent, according to him, smell) is perceptible, compared to the motion of light, which is phenomenologically inaccessible and hence appears immediate and static. In line with the phenomenological accessibility of the dynamic nature of sound, Pasnau (2000: 31), in his review of historical philosophical positions of sound notes that “it was obvious to the medievals that sound is closely connected with motion, perhaps identical to a certain kind of motion”. He furthermore notes that it took sophisticated technologies to uncover that sensible properties such as color and heat involve motion, “the same can be seen in the case of sound through intelligent observation” (ibid. 31), such as when seeing objects or water vibrate as a result of loud noises, or such as when feeling a direct blow through particularly loud noises.

As a point in comparison, consider the fact that taste and smell may be temporally variable as well: one may slowly become aware of a taste sensation, which grows stronger in one’s mouth and ultimately fades away, or one may notice a smell slowly unfolding in a room (cf. discussion in Huumo, 2010). However, both of these experiences lack a lot of the dynamicity that characterizes sound. The “dynamicity” of smell is relatively slow and furthermore mostly involves a rise and fall in intensity, compared to the internal temporal variation of the quality of sounds. Moreover, taste and smell may be more consistently associated with entities. We think of a particular entity as having the property of a particular taste or smell, the same way we think of a particular entity as having the property of being rough or smooth. Such properties, in contrast to sound, are seen as characteristic of the object without any form of noticeable action or movement. In sum, we are arguing that adjectives such as *sweet*, *stinky* and *rough* are understood as properties of an entity, compared to words such as *beeping* and *squealing*, which are understood primarily as actions of an entity.

The idea that sounds are more likely to be encoded via verbs as opposed to the other lexical categories has already been expressed by philosophers and linguists. For example, Pasnau (1999: 310) says that objects “do not have sounds, standardly, but instead make sounds. To squeak, squeal, wail, howl, quack — these are all ways of making a sound”. Krifka (2010) already observed (without quantitative evidence) that in English, sounds are commonly expressed via verbs and if they are expressed via other lexical categories, these are often deverbal (as with the noun *bang* and the adjective *banging*). Cognitive linguists have furthermore observed that sound may participate in WHOLE EVENT STANDS FOR COPRESENT SUBEVENT metonymies, as when saying *The car screeched to a halt* (Günter Radden, personal communication), where within a whole motion event the sound is metonymically profiled.

These observations by linguists and philosophers, in conjunction with our discussion of the phenomenology of sound, lead us to predict that verbs should be the favourite linguistic means to express sound and touch concepts. In what follows, we will provide a picture of the lexical category composition of English sensory lexicon that, among other things, will allow testing this hypothesis. As a final test and an additional dataset that substantiates the idea that sound concepts are very dynamic, we correlate movement and sound ratings for a large set of nouns from the Wisconsin perceptual attributes database (Medler et al., 2005). This dataset allows us directly assessing the question: are sound concepts more dynamic in the minds of English speakers? Altogether, our study not only addresses the question of whether the senses are lexicalized differently across lexical categories, but as far as we know, our study is the first to actually provide an empirical assessment of the idea that particular lexical categories have semantic preferences, in line with what cognitive and functional linguists have stated about the crucial involvement of time (Givón, 2001; Langacker, 2008).

3. Methods

3.1. Datasets

To investigate the distribution of sensory lexemes across lexical categories, we compared three different English datasets that include verbs, adjectives, and nouns. The datasets also provide information on the association of each lexeme with sensory modalities, which are in all three cases the five "classical" senses: sight, hearing, touch, taste, smell. Assuming five senses is non-trivial since it is philosophically unresolved how many distinct senses there are (see, for example, Macpherson, 2011; Casati et al., 2015), and since the senses are connected in highly complex ways (see, e.g., Spence & Piqueras-Fiszman, 2014). However, as a starting point it is useful to consider the five senses, which is also a way of categorizing the perceptual world that is adopted by the speakers of the language we look at, English (for discussion of five senses folk models in relation to methodological problems in sensory linguistics, see Winter, 2016b: Ch. 1).

Most work on sensory words has proceeded treating the problem of sensory modality classification as an easy one that can be solved in an intuitive fashion. For example, a lot of work on perception verbs, including Viberg (1983), Sweetser (1990) and San Roque et al. (2015), classified the verbs *to see* and *to look at* as visual, in contrast to the verb *to smell* and *to taste,* which would be classified as olfactory and gustatory respectively. In clear-cut cases such as these basic perception verbs, such a classification is straightforward. However, what about multisensory perception verbs such as *to perceive* (one can perceive a sound, a moving image but also a taste or smell)? Modality classifications are also difficult in the domain of adjectives, where there are some clear-cut unisensory adjectives such as *purple* (vision) and *beeping* (sound), but also many multisensory cases such as *harsh* (*harsh sound, harsh taste, harsh smell*) or magnitude or dimension terms such as *large* (Williams, 1976; cf. discussion in Ronga et al., 2012), which describe properties that can be perceived through multiple sensory modalities, what Aristotle and others since then call "common sensibles", percepts that are accessible to not just one but many senses (see Marks, 1978: Ch. 3). The problem of common sensibles is exacerbated when considering nouns, many of which can be perceived through all of the senses (see discussion in Lynott & Connell, 2013). For example, the concrete noun *steak* describes something that can be seen, felt, tasted, smelled and under the right conditions also heard. When doing quantitative comparisons across sensory modalities for many words, it is important to have some objective measure of sensory modality association. We need a defined set of words that has labels for sight, hearing, touch, taste and smell labels that have been assigned following clear criteria. Here, three such datasets will be considered.

As we will discuss below, each word list is associated with its own set of problems. This is precisely why it is important to consult multiple datasets, to be able to show that whatever results we find is truly generalizable, no matter the particular design decisions that went into collecting each word list. In fact, there can be no perfect list that meets all desiderata; each dataset is more or less compliant with certain theoretical assumptions. Because of this, we need to adhere to the principle of converging evidence (Lakoff & Johnson, 1999: 79-80; Gries, Hampe, Schönefeld, 2005), showing that a theoretical concept is supported by multiple sources and multiple different analyses.

Dataset 1 is a combined set of modality ratings of 423 adjectives from Lynott and Connell (2009), 400 nouns from Lynott and Connell (2013) and 300 verbs from Winter (2016a). Combining the three studies (adjectives, nouns and verbs), the word list from the norm dataset includes 1,123 words. The three studies adopted the same methodology: English native speakers were asked to rate whether a word corresponds to a particular sensory modality on a scale from 0 to 5. For example, speakers were asked to rate whether the property described by the adjective *fragrant* can be perceived through vision, hearing, touch, taste or smell, with five sliders embodying the relative "perceptual strength" of this property for the different sensory modalities. Following Lynott and Connell (2009), the maximum perceptual strength value can be taken to be a word's "dominant modality". For example, the adjective *shiny* received the highest mean rating for the visual modality, and it can hence be classified as a visual word.

The sampling procedures involved in the three studies were slightly different. The adjectives from Lynott and Connell (2009) are a convenience sample selected based on prior research (in particular, experimental studies that investigated sensory words) and based on synonyms and lexical field information gleaned from thesaurus lists. The noun list from Lynott and Connell (2013) is a random sample. The verbs from Winter (2016a) were comprised of two sets: one sample that was selected based on synonyms with basic perception verbs gleaned from the literature, particularly, Viberg (1983). And another sample that was selected randomly from words above median frequency in the English Lexicon Project (Balota et al., 2007). Although it would be desirable to have random samples for all three datasets, the results obtained in Winter (2016a) and Winter (2016b), for example, do not hinge on whether the random or the non-random subset of the verb is used—suggesting that to some extent the sampling procedure is not a big issue in this particular case. Moreover, in all of the three different studies the lists were not chosen with respect to the particular research question investigated in this paper, and the methods were principled (e.g., dictionary searches), suggesting that the potential role of bias in the sampling procedure specifically with respect to our research hypotheses is small. Moreover, precisely because of concerns about the sampling procedure in these three studies it is useful to show that we obtain qualitatively similar results for two other lists of sensory words.

Dataset 2 is a word list of sensory lexemes collected by Strik Lievers (2015). This list was collected starting from a short list of core sensory lexemes, which was expanded in successive phases using various lexical resources and dictionaries. For instance, the noun list was enriched by searching for the direct objects that display stronger association with the perception verbs already in the list, based on corpus data; adjectives were retrieved among the modifiers of perception nouns; synonyms and hyponyms were obtained for all lexemes. At each step the list was manually checked (see Strik Lievers, 2015: 76-78 and Strik Lievers & Huang, 2016 for details).

The dataset was created to retrieve instances of synesthetic metaphors from corpus data. The list was subsequently extended to include a total of 512 words (Strik Lievers & Huang, 2016), from which all nouns of musical instruments (*trumpet*, *piano*, etc.) were excluded (these were previously classified as auditory). The total dataset of words without instruments comprises 486 words.

Dataset 3 is the Sensicon (Tekiroğlu et al., 2014), a resource which includes 22,684 lexemes together with their degree of association with the five senses. A second list has subsequently been published by the same authors (Tekiroğlu et al., 2015), however, this list does not include verbs and thus cannot be used to address the present research hypotheses. The Sensicon dataset we use here was constructed by looking at word co-occurrence statistics in the GigaWord corpus. The idea here is that if a given word occurs very frequently together in the same text with a particular seed word of a given sensory modality, then it is a word of this sensory modality (for problems with this assumption, see Louwerse & Connell, 2011, Winter, 2016b, and Tekiroğlu et al., 2015). The structure of the dataset is similar to the ratings represented in Lynott and Connell (2009, 2013) and Winter (2016a), with a continuous numerical value for each modality association. Because the classification in Strik Lievers (2015) is discrete (a word is either of a given sensory modality or not) and because the sensory modality ratings by Lynott and Connell (2009, 2013) and Winter (2016a, 2016b) have been used in a discrete fashion, which facilitates computations of lexical differentiation, we will also treat the Sensicon ratings in a discrete fashion, selecting a word's highest numerical value as that word's dominant sensory modality.

As mentioned before, the classifications in each dataset are not entirely unproblematic, and there is considerable noise associated with some of the datasets. For example, Strik Lievers (2015) did not consider *ticklish, painful* and *tingly* as touch words, but they are counted as touch words in the Lynott and Connell (2009) dataset because the highest ratings of these words were for the tactile modality. Many of the nouns in the Lynott and Connell (2013) dataset are highly abstract (e.g., *heaven, fact*) and not at all strongly related to sensory perception, compared to the Strik Lievers (2015) nouns, which more directly relate to perception (e.g., *glare, rustle, gleam, shadow, tune*). Lynott and Connell (2013) discuss the abstractness of many of the nouns involved, which is a natural outgrowth of the fact that these words were randomly sampled. In particular, in their dataset many abstract nouns were rated to be highest in auditory content (such as *account, blame*), presumably because participants thought that they were mediated through language and speech, and thus sound-related in some way. Similarly, because they are based on text cooccurrences, it is not at all clear in many cases what the modality associations for the Sensicon mean, with many examples that just seem “off”, such as *inspector* being classified as auditory, and *fraction* as olfactory.

3.2. Analysis approach

How should we cope with the fact that we are dealing both with proper perception concepts, that is words that can adequately be called “sound words”, “touch words” etc., as well as perception-*related* concepts, and ultimately words that are so highly abstract (e.g., *heaven*) or multimodal (e.g., *seem*) that classifying them according to sensory perceptions makes no sense? Luckily, there are multiple ways of dealing with this problem. For the norm dataset and the Sensicon, we can use the continuous measures of modality associations to get a “cleaner” dataset of words that are more strongly related to actual perception. First, Lynott and Connell (2009) compute a measure of “modality exclusivity”, ranging from 0% (all five senses the same) to 100% (no overlap in ratings for the five senses). For example, the adjective *blue* had a modality exclusivity of 80%, indicating that it was highly visual in a unisensory fashion. The adjective *strange* on the other hand had a modality exclusivity of 9.6%, indicating that it does not relate very strongly to any particular sense. In one of our analyses, we included only words that were above the 70th percentile on this modality exclusivity measure; that is, we excluded highly multimodal words. As a second exclusion criterion, we considered overall perceptual strength, that is, the sum of the five modality ratings (see also Connell & Lynott, 2012). For example, the highly non-sensory words *republic* (+2.79), *remark* (+2.9) and *corrupt* (+3.33) had very low perceptual strength ratings overall, compared to the much more sensory words *silky* (+9.29), *short* (+9.04) and *bitter* (+8.95). We ran an additional set of analyses with only those words that were above the 70th percentile in the overall perceptual strength measure. These two exclusion criteria can be applied to both the modality norm datasets, as well as to the Sensicon, because both have continuous perceptual strength measures. As a third exclusion criterion, one of the authors (Strik Lievers) marked words in the norm dataset as questionable with respect to whether they had any sensory qualities at all. We have in front of us what Gelman and Loken (2013) call the “garden of forking paths” when performing a statistical analysis, which is potentially dangerous because it invites researcher degrees of freedom (Simmons et al., 2011). Rather than ignoring these potentially problematic analysis decisions, we make them an integral part of our analysis. The R script (using R version 3.3.1, R Core Team, 2015) that we make available with this publication has several “switches” for running the analysis with (1) only the 30% most unimodal words, (2) only the 30%

words with the highest perceptual strength and (3) only those words that were not flagged as questionable. As a fourth and final “switch” in the analysis, we can consider lexical category counts only for those words that are above median frequency in SUBTLEX, the Subtitle Corpus of American English (Brysbaert & New, 2009). The main findings we report below, the “verbiness” of sound concepts, can be found in all three datasets under all combinations of these “analysis switches”.

We thus assess the sensitivity of our findings with respect to several analysis decisions, finding that the main result holds. Moreover, as is clear from the brief descriptions above, the three datasets have been built independently from each other, for different research purposes and with different methods. Rather than being a problem, this is an advantage—to the extent that we show results that hold across these three different datasets, these results are supported by converging evidence and are thus more generalizable. If the datasets were all constructed using the same sampling criteria or the same approach to sensory classification, our results would be less convincing because they may be subject to these particular methodological decisions. By using three different datasets, we circumvent this concern.

Below, we report analyses for all three datasets, for convenience sake only showing the results for (1) the highly unimodal ones (70th percentile exclusivity), (2) all words regardless of perceptual strength (no exclusion), (3) all frequencies (no exclusion) and (4) without the words that were flagged as questionable. While the results, of course, differ in terms of the precise numerical values if different exclusion criteria are used, the substantive conclusions do not change.

As a final methodological decision, we need to talk about how lexical category assignments have been dealt with. Each of the three word lists comes with its own set of lexical category labels. The main result can be established with the existing category labels that come with each dataset; however, there are several problems. Consider the word *squealing*, which is part of the Lynott and Connell (2009) adjective ratings, but could also be labeled as verb based on its morphology (the suffix *–ing*). However, morphology is not a good criterion in all cases (and we would need to take productivity of affixes into account), and moreover it does not help in cases of zero derivation (e.g., *blue* can be used as adjective and noun). To deal with such decisions in a principled manner, we used corpus-based lexical category classifications from Brysbaert et al. (2012). These researchers used an automatic tagger on the SUBTLEX subtitle corpus to determine whether a word was used as a noun, verb or adjective. For example, the word form *squealing* occurs 97 times in SUBTLEX as a verb and only 5 times as an adjective. We used the corpus-based tags from Brysbaert et al. (2012), only including those words that occurred 70% in their primary lexical category.

We should note that in using corpus-based part-of-speech tags, we are explicitly assuming that nouns/verbs/adjectives are distributionally defined, and we then look for semantic differences across these distributionally defined categories (cf. Baker & Croft, 2017: 183), in this case in terms of perceptual meaning. Thus, we are essentially testing semantics within a predefined set of lexicogrammatical differences.

The scripts are accessible together with the data under the following publically accessible repository[2]:

[omitted because of requirement to be anonymous]

4. Results

4.1. Lexical category counts

We start out by considering the differentiation of lexical categories across the five sensory modalities, essentially just performing a type count, i.e., how many unique words are there per sensory modality per lexical category. We cannot, however, simply count up sight words, sound words, touch words, etc. per lexical category etc. This is because in each dataset, there is a different baseline number of lexical categories: the modality ratings have 36% nouns, 38% adjectives and 27% verbs; Sensicon has 51% nouns, 31% adjectives and 17% verbs; the Strik Lievers (2015) dataset has 40% nouns, 36% adjectives and 24% verbs (before any of the exclusion criteria are applied). This means that we have to take into account that there are less verbs overall. In addition, we have to take into account that certain sensory modalities have more or less word types associated with them, regardless of lexical category. For example, the datasets exhibit show some form of visual bias (see Levinson & Majid, 2014; Strik Lievers, 2015; Winter, 2016b): in the modality ratings, there are 57% visual words (followed by 16% auditory, 16% touch, 7% taste, 4% smell). In the Sensicon, there are 27% visual words (followed by 26% taste, 19% touch, 18% sound, 10% smell). In the Strik Lievers (2015) dataset, there are 45% sound words (followed by 29% sight, 10% touch, 9% taste, 6% smell). The tables in Appendix A list the counts per sense and per lexical category for all three datasets (no exclusions).

[2] All analyses were conducted with the packages “tidyverse” version 1.1.1 (Wickham, 2017), “stringr” version 1.1.0 (Wickham, 2016) and “png” version 0.1-7 (Urbanek, 2013). We use the following sense logos from FreePik: http://www.freepik.com/free-vector/five-senses-icons_837465.htm

Thus, in terms of raw word type counts per sensory modality, there are stark differences between the three datasets, but also some clear similarities. First of all, vision is ranked first in two out of three cases, supporting the idea that the English lexicon exhibits “visual dominance” (Buck, 1949; Viberg, 1983; Levinson & Majid, 2014; Winter, 2016b: Ch. 3). That is, the English language appears to make more distinctions in the visual modality than in any other sensory modality. A second result that is consistent across all three datasets is that smell is always ranked last, vindicating the view that at least in English, there are very few words for smells (Buck, 1949; Viberg, 1983; Levinson & Majid, 2014; Winter, 2016b: Ch. 3) and the more general idea that smell is a “muted sense” (Yeshurun & Sobel, 2010; Olofsson & Gottfried, 2015). With the exception of the Sensicon, where taste ranks very highly, vision, touch, and sound appear to rank together as having high word counts, followed by taste and smell.

If we now want to look at whether particular lexical categories are over- or under-represented for a particular sense, we need to keep these asymmetries between the senses in mind, together with any asymmetries between overall lexical category counts. To do this in a principled manner, we use Chi-Square tests and standardized Pearson residuals. The Chi-Square test computes the expected counts for each cell, based on a simple multiplication of the row total (how many words per lexical category) and the column total (how many words per sensory modality). Standardized Pearson residuals then give a standardized measure of how much each unique cell in a cross-tabulation deviates from expected counts.

In the following, we will report data based on the “cleaned” set with the exclusion criteria stated above (70th percentile unimodal, excluding questionable cases). In the case of the modality norm datasets (with N = 196 after exclusion), a Chi-Square test yields a significant result (χ^2 = 43.25, bootstrapped *p-value* based on 2,000 samples p = 0.00005). This provides a formal test of the idea that indeed, lexical categories are not distributed evenly across the five senses. Figure 1 displays the standardized residuals, with values larger than 2 or smaller than -2 being indicative of contributing to a significant Chi-Square value (see Levshina, 2015: 220-221; Agresti, 2001). As can be seen, based on this >|2| cut-off rule, verbs are over-represented particularly for sound and touch. They are significantly under-represented for vision, which has comparatively more nouns and adjectives than what is expected based on chance. In the modality norm dataset, touch has significantly less adjectives, and sound has significantly less nouns.

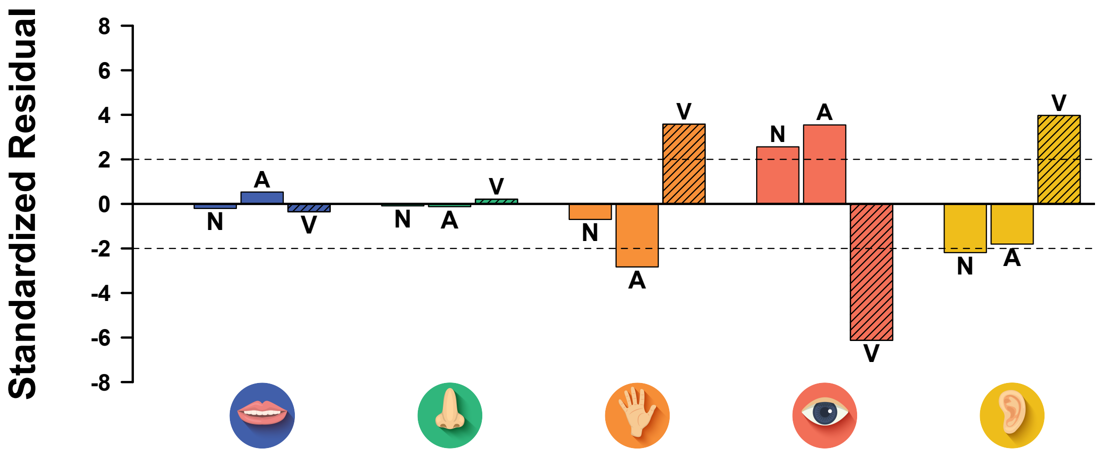

Figure 1: Standardized Pearson residuals for the three lexical categories per sensory modality; based on data from Lynott & Connell (2009, 2013) and Winter (2016a)

Besides this result that looks across lexical categories, it should be noted that even only within the list of adjectives by Lynott and Connell (2009), there is a clear verb bias: of the 68 properties, 42 (62%) turned out to be more frequently used as verbs when the corpus-based part-of-speech tags from Brysbaert et al. (2012) were used. Not only does hearing have many more verbs than nouns or adjectives, but also many of those words that are treated as auditory adjectives in the psycholinguistic literature that is based on the Lynott and Connell (2009) ratings are actually deverbal. Semantically, this means that many auditory adjectives (such as *squealing*) are not prototypical property-describing adjectives: they rather describe actions, thus contributing to increase the semantic "verbiness" of the auditory lexicon.

For the Strik Lievers (2015) dataset, there were 254 unique words left after exclusion. Again, lexical categories were distributed unevenly across the five senses (χ^2 = 38.0, p = 0.0005). A look at the Pearson residuals in Figure 2 reveals that there are many differences with respect to the previous dataset with respect to which lexical categories are over- or under-represented for which sensory modalities. In particular, whereas adjectives were significantly under-represented in the norm dataset for touch, they are significantly over-represented for touch in the Strik Lievers (2015) dataset. Touch also has significantly less nouns and verbs in this dataset. Taste has significantly more adjectives. For sound, there were significantly more nouns and significantly less adjectives. Crucially, despite all these discrepancies to the previous dataset, verbs are still over-represented for sound.

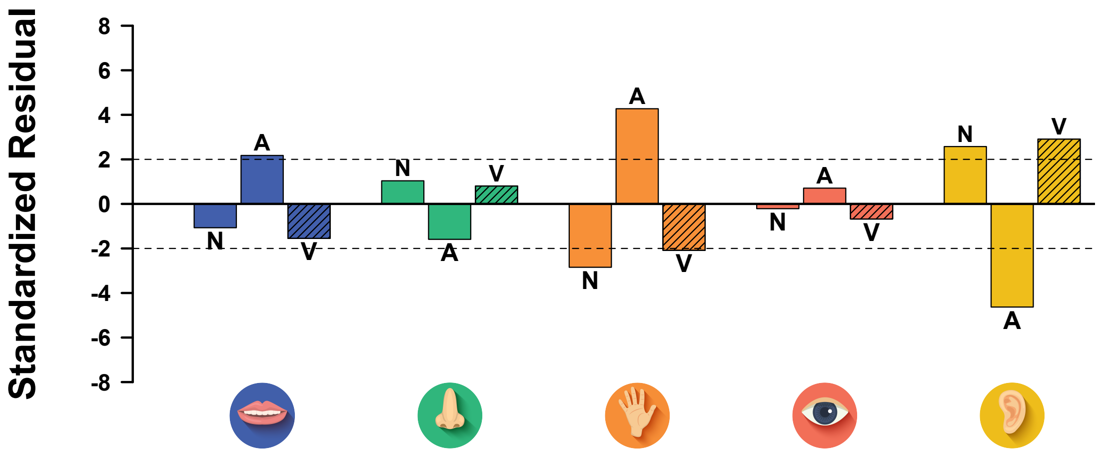

Figure 2: Standardized Pearson residuals for the three lexical categories per sensory modality; based on data from Strik Lievers (2015)

Finally, what about the Sensicon (N = 4,405 after exclusion)? Again, lexical categories were unevenly distributed across the five senses (χ^2 = 102.9, p = 0.0005). A look at the Pearson residuals in Figure 3 reveals that adjectives were significantly over-represented for smell and sight, and under-represented for touch and sound. Nouns were significantly under-represented for sight. Finally, verbs were significantly over-represented for touch, under-represented for sight, and crucially, over-represented for sound.

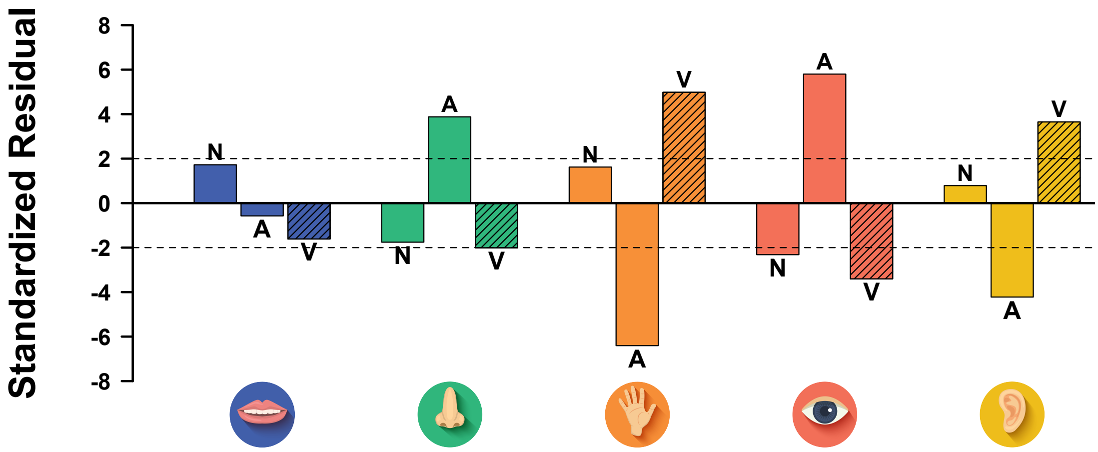

Figure 3: Standardized Pearson residuals for the three lexical categories per sensory modality; based on data from the Sensicon (Tekiroğlu et al., 2014)

To look at all datasets together, we average the standardized residuals for each cell across the ratings, the Strik Lievers (2015) word list and the Sensicon. This revealed that verbs were over-represented for sound (+3.5) and adjectives were under-represented (-3.5). On top of that, adjectives were over-represented for sight (+3.3) and verbs were under-represented for sight (-3.0). No other Pearson residuals crossed the threshold of |2|. The over-representation of adjectives for sight, compared to the under-represented for sound, is noteworthy. Sound and sight are traditionally regarded as being on top of a "hierarchy" of the senses (e.g., Ullmann, 1959 [1957]), yet the auditory modality has less words that describe dedicated perceptual characteristics in an adjectival fashion.

4.2. Wisconsin perceptual attribute ratings

We now move away from considerations of lexical category differences and provide one piece of data that lends additional support for the dynamic nature of sound concepts that stems from an independent dataset and a different approach. The dataset we consider here, the Wisconsin perceptual attributes database, is a set of 1,402 words that have been rated for how much they make reference to particular domains of sensorymotor experience (the description of the ratings does not discuss what criteria were used to sample the word list). 342 undergraduate students rated different semantic dimensions on a scale from 0 (concept not at all important for this dimension) to +6 (very important). Crucially, there are three semantic dimensions in this dataset relevant to our idea of the dynamicity of sound, namely: sound ratings, motion ratings and, for comparison, color ratings. The words that received the highest sound ratings were *explosion, siren, scream, bomb, fireworks, dynamite, rocket, gunshot, thunder* and *alarm*. The words with the lowest sound ratings were *palm, prune, velvet, broccoli, oblique, sum, yam, mushroom, corpse* and *number*. The words that received the highest color ratings were *orange, bluejay, blueberry, flamingo, rose, tomato, pumpkin, cherry, sun,* and *autumn.* The words that received the lowest color ratings were *actuality, heresy, interim, lecture, remedy, reprisal, agility, analogy, aye,* and *bequest*. Finally, the words with the highest motion ratings were *rocket, tornado, cheetah, hurricane, jet, avalanche, stampede, sex, jaguar,* and *children*. The words with the lowest motion ratings were *asphalt, basement, box, bread, brick, cabin, cave, ceiling, cliff* and *corn.* Given the idea that sound is more dynamic, we would expect a higher correlation between sound and motion ratings than between motion and color ratings. Here, we operationalize dynamicity with respect to motion ratings alone (as outlined above, the concept of dynamicity in the domain of sound is wider than just the involvement of movement).

Pairwise correlations, depicted in Figure 4, show that sound ratings are reliably correlated with motion ratings ($t(1400) = 27.28$, $p < 0.0001$), with a relatively high correlation coefficient, Pearson's $r = 0.59$ ($R^2 = 0.35$). There also was a reliable correlation with color ratings ($t(1400) = 8.03$, $p < 0.0001$), but a much smaller one with $r = 0.21$ ($R^2 = 0.04$). A simple linear regression model where both sound and color ratings are used to predict motion ratings with an interaction for type of rating (color versus sound) shows that the slope for sound and motion is reliably stronger than the slope for color and motion (estimate: 0.32, $SE = 0.03$, $t = 11.06$, $p < 0.001$).

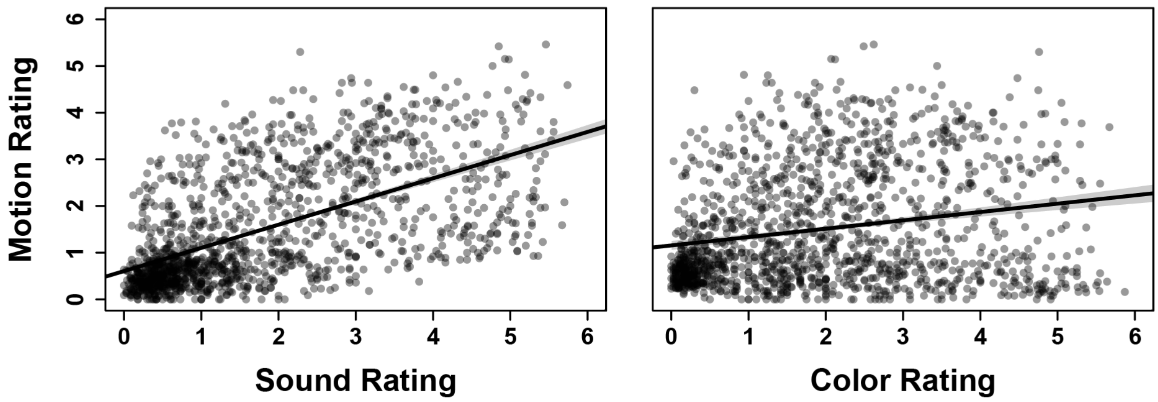

Figure 4: Correlation between sound/color ratings with motion ratings; superimposed fit of a simple linear regression with 95% confidence region

These results provide independent evidence for the idea that sound-related concepts tend to also be motion-related concepts. These ratings were performed on words from one lexical category alone (nouns), but even within that lexical category, there is evidence for sound-related concepts being thought of as relatively more dynamic, at least when compared to color-related concepts. Given the limits of the Wisconsin ratings (which did not include ratings for the other sensory modalities), here, only comparisons between sound and color were possible, for which sound emerged as more motion-related than color. However, this evidence corroborates what we found for the sound-verb association when looking at lexical category differences across the senses, except that in this case, participants directly rated the *meaning* of the words involved with respect to motion, sound and color.

5. Discussion

The comparative analysis of four datasets, each collected independently, showed that the composition of the English sensory lexicon is not uniform across sensory modalities, and that sound-related concepts are associated with dynamicity. For the lexical category results, we showed that regardless of which dataset was consulted, and even though the datasets differed quite starkly with respect to the ranking of the other senses, verbs were always over-represented in the domain of sound. We furthermore found corroborating evidence for the idea of visual dominance, as well as for the idea that smell is lexically impoverished in English. As concerns the overall number of words per sensory modality, vision ranks first in two out of three datasets (followed by hearing and touch, and in the case of the Sensicon, followed by taste). Smell consistently ranked last. This provides quantitative confirmation of the common depiction of vision as the dominant sense and olfaction as a “muted” sense in the English lexicon (cf. Levinson & Majid, 2014). Second, as concerns the distribution of lexical categories across the senses, the analysis showed that verbs are overrepresented for hearing. This was the case for all three different datasets. In two out of three datasets, we also found that verbs were over-represented for touch.

The unequal distribution of lexical categories across sensory modalities turned out to be consistent on the one hand with the semantic properties of prototypical members of the relevant lexical category, and on the other hand with the properties of each of the five senses in actual perception, as indicated by our review of the literature of the phenomenology of auditory perception. In particular, the results of our analysis confirmed the hypothesis that, given that prototypical verbs describe actions, events and processes, and given that sound and (to a minor extent) touch are highly dynamic sensory modalities, verbs are particularly fit to express auditory and tactile experiences. The connection between hearing and verbs seems to be stronger than the connection between the other senses and the verbal domain.

It should be noted that all of these patterns are probabilistic. There clearly are verbs associated with each sensory modality (e.g., the basic perception verbs *to see, to hear, to feel, to smell*, *to taste*), and there are also adjectives associated with each sensory modality (e.g., *purple, loud, rough, musky, bitter*), and also nouns (e.g., *image, melody, contact, odor, flavor*). The patterns we discuss here are not all-or-nothing, but they are about the relative degree to which particular senses tend to associate with particular lexical categories.

The patterns we found here fit with existing literature on language and the senses in cognitive linguistics, or functional-cognitive linguistics more generally. Huumo (2010) provided independent evidence for sound being more dynamic in his analysis of which locative markers go together with which perception verbs in Finnish. He observed that active perception verbs tend to have directional case markers (‘from’ or ‘to’), and this also characterized sound verbs (and smell verbs), but not visual verbs, which tended to go with more “static” case markers (‘in’, ‘on’, ‘at’). Similarly, the dynamicity of touch has been noted by Popova (2005), although her analysis focused not on verbs but on the gradability in adjectives. Popova (2005: 400) described touch as an “active sense”, stating that “the most common mode of touch is the active movement of the hand”. She furthermore cites Katz (1989 [1925]: 242) who said that “[t]ouching means to bring to life a particular class of physical properties through our own activity.” These arguments are in line with our finding that touch, although not as much as sound, latches onto the verbal domain relatively more strongly. Finally, Winter, Perlman, Perry and Lupyan (2017) argue that the dynamicity of both touch and sound as sensory modalities may explain why both sound and touch words in English are so highly “onomatopoetic” in character, since vocal iconicity may be particularly effective if a dynamic medium (sound and touch) is expressed in another dynamic medium (speech) (see also Winter, 2016b: Ch. 6).

The distribution of lexical categories with respect to the senses is also relevant for studies of synesthetic metaphors. Ullmann (1959 [1957]) already remarked that asymmetries in the vocabularies of languages could lead to asymmetries in metaphors, i.e., senses that have less lexical material associated with them need to “borrow” more words from the other senses, an argument that was extended by Strik Lievers (2015) to be specifically about lexical categories. Our results provide further evidence for this claim. In particular, they may contribute to account for the fact that in the literature on synesthetic metaphors sound — rather than the “dominant” modality of sight — consistently emerges as the ultimate target domain of cross-sensory mappings, often with sight as the source sensory modality (Ullmann, 1959 [1957]; Williams, 1976; Day, 1996; Shen, 1997; Strik Lievers, 2015; Winter, 2016b). The three datasets show that adjectives are under-represented for sound, while they are over-represented for sight. Taking into consideration that the dominant pattern investigated in the literature on synaesthesia are adjective-noun combinations, the adjective being the source and the noun the target domain, it is therefore not surprising that, between sight and sound, it is sight that is more commonly found as a source. Given this, some asymmetries observed in cross-sensory mappings may correlate with lexical category differences (as suggested by Strik Lievers, 2015; other factors have to be taken into account as well, see Winter, 2016b: Ch. 8). Either way, our results suggest that when people perform metaphor counts, they should take into account what an appropriate “baseline” for comparison is and what specific affordances are created by the lexicon.

What do our results say about research on lexical categories? As stated by Baker and Croft (2017: 181), “Prior to the advent of structuralist notions of categories, the widespread view was that lexical categories were defined notionally, by something like the idea that nouns express things, verbs express actions, and adjectives express properties.” We do not say that we have to go back to a fully notional view of lexical categories, especially since morphological and distributional criteria (especially if amended by other criteria, see Farmer et al., 2006) work so well within language. However, we should point out that besides the introspective analyses in works such as Givón (1979), Givón (2001 [1984]) and Langacker (2008), there was, so far, little *quantitative* evidence for systematic semantic differences between different lexical categories. The idea that lexical categories defined by their grammatical properties differ in meaning has so far been a claim that only rested on intuitions; here, we tested this general idea using the specialized vocabulary of sensory language. Of course, sensory meaning does not exhaust semantics; it is only a narrow subpart of it. Nevertheless, in this case, sensory meaning proved to be a useful access point for investigating semantic differences between lexical categories. We hope that further studies of lexical categories will incorporate similar methodologies, such as the use of human rating studies such as Lynott and Connell (2009) to quantify the semantic intuitions and the claims that have been put forth in cognitive linguistics.

As reviewed above, already Aristotle viewed nouns as having “no reference to time” and sounds as not being able to occur without movement. In this paper, we connect these two claims. We should note, however, that whereas Aristotle thought of this in terms of actual ontologies, here we are talking more about conceptualization. Baker and Croft (2017: 118) state that “The semantic contrast in the linguistic expressions, including the lexical category that is used, reflects that conceptualization, not the “objective” properties of the entities being described.” However, the way sound is produced “objectively” in the world, namely through a dynamic event, fosters a consistent phenomenology of sound as a time-varying and motion-related quality. This conceptualization, in turn, may drive how sound is encoded in the lexicon, such as the present evidence from English showed. All senses involve motion and action to some extent, but in the case of sound this is phenomenologically more apparent to the language users, which hence may drive particular forms of linguistic encoding.

Finally, we hope to have shown on the methodological side that many interesting questions can be asked, and answered, by using already existing datasets. In our case, we used humanly generated property ratings (Lynott & Connell, 2009), noun ratings (Lynott & Connell, 2013), verb ratings (Winter, 2016a), sound and motion ratings (Medler et al., 2005), a manually annotated lexicon (Strik Lievers, 2015) and the automatically generated Sensicon (Tekiroğlu et al., 2014) to address questions about language and perception, as well as about the semantics of lexical categories more generally. While there was a lot of noise in the used data sources, applying the principle of converging evidence (Lakoff & Johnson, 1999: 79-80; Gries, Hampe, Schönefeld, 2005) through the incorporation of multiple data sources allowed us to draw confident conclusions.

To conclude: not all senses are created equal. While the senses may be differentially encoded in general (Levinson & Majid, 2014), the present data demonstrated that the senses may also be differentially encoded with respect to lexical categories in particular. The senses latch onto particular domains of experience, and depending on what type of experiences are preferentially expressed through certain lexical categories, such as verbs being more dynamic, this creates asymmetries in how perception is encoded in language.

References

Aarts, B., & Haegeman, L. (2006). English word classes and phrases. In B. Aarts and A. McMahon (Eds.), *The Handbook of English Linguistics* (pp. 116-145). Malden, MA: Blackwell Publishing.

Agresti, A. (2002). *Categorical Data Analysis* (2nd Edition). Hoboken, New Jersey: John-Wiley & Sons.

Baker, M. (2003). *Lexical categories: Verbs, nouns, and adjectives*. Cambridge: Cambridge University Press.

Baker, M., & Croft, W. (2017). Lexical categories: Legacy, lacuna, and opportunity for functionalists and formalists. *Annual Review of Linguistics*, *3*, 179-197.

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, *39*, 445-459.

Bartley, S. H. (1953). The perception of size or distance based on tactile and kinesthetic data. *The Journal of Psychology*, *36*, 401-408.

Berg, T. (2000). The position of adjectives on the noun-verb continuum. *English Language and Linguistics*, *4*, 269-293.

Blevins, J. P. (2012). Word-based morphology from Aristotle to modern WP. In K. Allan (Ed.), *The Oxford Handbook of the History of Linguistics* (pp. 397-417). Oxford: Oxford University Press.

Bod, D. (2015). Probabilistic linguistics. In B. Heine and H. Narrog (Eds.), *Oxford Handbook of Linguistic Analysis* (pp. 663-691). Oxford: Oxford University Press.

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*, 977-990.

Brysbaert, M., New, B., & Keuleers, E. (2012). Adding part-of-speech information to the SUBTLEX-US word frequencies. *Behavior Research Methods*, *44*, 991-997.

Buck, C. D. (1949). *A Dictionary of Selected Synonyms in the Principal Indo-European Languages: A Contribution to the History of Ideas*. Chicago: University of Chicago Press.

Carlson, N. R. (2010). *Physiology of Behavior* (10th Edition). Boston: Allyn & Bacon.

Casati, R., Dokic, J., & Le Corre, F. (2015). Distinguishing the commonsense senses. In D. Stokes, M. Matthen & S. Biggs (Eds.), *Perception and its Modalities* (pp. 462-479). Oxford: Oxford University Press.

Cassidy, K. W., & Kelly, M. H. (1991). Phonological information for grammatical category assignments. *Journal of Memory and Language*, *30*(3), 348-369.

Classen, C. (1993). *Worlds of Sense: Exploring the senses in history and across cultures*. London: Routledge.

Connell, L., & Lynott, D. (2012). Strength of perceptual experience predicts word processing performance better than concreteness or imageability. *Cognition*, *125*(3), 452-465.

Croft, W. (1991). *Syntactic Categories and Grammatical Relations: The Cognitive Organization of Information*. Chicago: University of Chicago Press.

Day, S. (1996). Synaesthesia and synaesthetic metaphors. *Psyche*, *2*, 1-16.

Evans, N. M. & Wilkins, D. (2000). In the mind's ear: The semantic extension of perception verbs in Australian languages. *Language, 76*(3), 546-592.

Farmer, T. A., Christiansen, M. H., & Monaghan, P. (2006). Phonological typicality influences on-line sentence comprehension. *Proceedings of the National Academy of Sciences*, *103*(32), 12203-12208.

Frawley, W. J. (1992). *Linguistic Semantics*. Hillsdale, NJ: Lawrence Erlbaum.

Gärdenfors, P. (2014). *The Geometry of Meaning: Semantics Based on Conceptual Spaces*. Cambridge, MA: MIT Press.

Gelman, A., & Loken, E. (2013). The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time. In Technical report, Department of Statistics, Columbia University. [www.stat.columbia.edu/~gelman/research/unpublished/p_hacking.pdf , Retrieved July 31, 2017]

Gibson, J. J. (1962). Observations on active touch. *Psychological Review*, *69*(6), 477– 90.

Givón, T. (1979). *On Understanding Grammar.* Orlando, FL: Academic Press.

Givón, T. (2001 [1984]). *Syntax: A Functional-Typological Introduction*, Vol. 1. Amsterdam-Philadelphia: Benjamins.

Gries, S. T., Hampe, B., & Schönefeld, D. (2005). Converging evidence: Bringing together experimental and corpus data on the association of verbs and constructions. *Cognitive Linguistics*, *16*, 635-676.

Hamlyn, D. W (1968). *Aristotle De Anima: Books II and III. Translated with Introduction and Notes*. Oxford: Clarendon Press.

Hartmann, W. M. (1995). The physical description of signals. In B. Moore (Ed.), *Hearing* (pp. 1- 40). San Diego: Academic Press.

Himmelmann, N. P. (2017). Word classes. In *Oxford Bibliographies in Linguistics*, http://www.oxfordbibliographies.com/view/document/obo-9780199772810/obo-9780199772810-0159.xml [accessed November 8, 2017]

Hollmann, W. B. (2013). Nouns and verbs in Cognitive Grammar: Where is the 'sound'evidence?. *Cognitive Linguistics*, *24*(2), 275-308.

Howes, D. (2003). *Sensual Relations: Engaging the Senses in Culture and Social Theory*. Ann Arbor, MI: University of Michigan Press.

Huumo, T. (2010). Is perception a directional relationship? On directionality and its motivation in Finnish expressions of sensory perception. *Linguistics*, *48*(1), 49-97.

Ježek, E. (2016). *The Lexicon. An Introduction*. Oxford: Oxford University Press.

Katz, D. (1989 [1925]). *The World of Touch*. Edited and translated by Lester E. Krueger. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Kelly, M. H. (1988). Phonological biases in grammatical category shifts. *Journal of Memory and Language*, *27*(4), 343-358.

Kelly, M. H. (1992). Using sound to solve syntactic problems: the role of phonology in grammatical category assignments. *Psychological Review*, *99*(2), 349-364.

Kelly, M. H., & Bock, J. K. (1988). Stress in time. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(3), 389-403.

Krifka, M. (2010). A note on an asymmetry in the hedonic implicatures of olfactory and gustatory terms. In S. Fuchs, P. Hoole, C. Mooshammer & M. Zygis (Eds.), *Between the Regular and the Particular in Speech and Language* (pp. 235-245). Frankfurt: Peter Lang.

Kulvicki, J. (2015). Sound stimulants: Defending the stable disposition view. In D. Stokes, M. Matthen & S. Biggs (Eds.), *Perception and its Modalities* (pp. 205-221). Oxford: Oxford University Press.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the Flesh*. New York: Basic books.

Langacker, R. W. (2008). *Cognitive Grammar: A Basic Introduction*. Oxford, UK: Oxford University Press.

Lehrer, A. (1978). Structures of the lexicon and transfer of meaning. *Lingua, 45*, 95-123.

Levinson, S. C., & Majid, A. (2014). Differential ineffability and the senses. *Mind & Language*, *29*, 407-427.

Levshina, N. (2015). *How to do linguistics with R: Data exploration and statistical analysis*. Amsterdam: John Benjamins.

Liberman, M., & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry*, *8*(2), 249-336.

Lynott, D., & Connell, L. (2009). Modality exclusivity norms for 423 object properties. *Behavior Research Methods*, *41*, 558-564.

Lynott, D., & Connell, L. (2013). Modality exclusivity norms for 400 nouns: The relationship between perceptual experience and surface word form. *Behavior Research Methods*, *45*, 516-526.

Lyons, J. (1977). *Semantics*. Cambridge: Cambridge University Press.

Macpherson, F. (Ed.) (2011). *The Senses: Classic and Contemporary Philosophical Perspectives*. Oxford: Oxford University Press.

Majid, A. (2015). Cultural factors shape olfactory language. *Trends in Cognitive Sciences*, *19*(11), 629-630.

Majid, A., & Burenhult, N. (2014). Odors are expressible in language, as long as you speak the right language. *Cognition*, *130*(2), 266-270.

Marín, R. (2010). Spanish adjectives within bounds. In P. C. Hofherr & O. Matushansky (Eds.), *Adjectives: Formal Analyses in Syntax and Semantics* (pp. 307-332). Amsterdam: John Benjamins.

Marks, L. E. (1978). *The Unity of the Senses: Interrelations Among the Modalities*. New York: Academic Press.

Maslova, E. (2004). A universal constraint on the sensory lexicon, or when hear can mean 'see'? In A. P. Volodin (Ed.), *Tipologičeskie obosnovanija v grammatike: k 70-letiju professora Xrakovskogo V.S.* (Typological Foundations in Grammar: For Professor Xrakovskij on the Occasion of his 70th Birthday), (pp. 300-312). Moscow: Znak.

Matthen, M. (2010). On the diversity of auditory objects. *Review of Philosophy and Psychology*, *1*(1), 63-89.

Medler, D. A., Arnoldussen, A., Binder, J. R., & Seidenberg, M. S. (2005). The Wisconsin perceptual attribute ratings database. [http://www.neuro.mcw.edu/ratings/, retrieved July 2, 2017]

Miller, G. A., & Johnson-Laird, P. N. (1976). *Language and Perception*. Cambridge: Harvard University Press.

Monaghan, P., Chater, N., & Christiansen, M. H. (2005). The differential role of phonological and distributional cues in grammatical categorisation. *Cognition*, *96*(2), 143-182.

Monaghan, P., Christiansen, M. H., & Chater, N. (2007). The phonological-distributional coherence hypothesis: Cross-linguistic evidence in language acquisition. *Cognitive Psychology*, *55*(4), 259-305.

Murphy, M. L. (2010). *Lexical Meaning*. Cambridge: Cambridge University Press.

Nudds, M. (2015). Audition. In M. Matthen (Ed.), *The Oxford Handbook of Philosophy of Perception* (pp. 274-293). Oxford: Oxford University Press.

O'Callaghan, C. (2007). *Sounds: A Philosophical Theory*. Oxford: Oxford University Press.

O'Callaghan, C. (2008). Object perception: Vision and audition. *Philosophy Compass*, *3*(4), 803-829.

O'Callaghan, C. (2009). Sounds and events. In M. Nudds & C. O'Callaghan (Eds.), *Sounds & Perception. New Philosophical Essays* (pp. 26-49). Oxford: Oxford University Press.

O'Callaghan, C. (2014). Auditory Perception. In *Stanford Encyclopedia of Philosophy*. https://seop.illc.uva.nl/entries/perception-auditory/ [accessed October 12, 2017]

O'Callaghan, C. & Nudds, M. (2009). Introduction: The philosophy of sounds and auditory perception. In M. Nudds & C. O'Callaghan (Eds.), *Sounds & Perception. New Philosophical Essays* (pp. 1-25). Oxford: Oxford University Press.

O'Meara, C., & Majid, A. (2016). How changing lifestyles impact Seri smellscapes and smell language. *Anthropological Linguistics*, *58*(2), 107-131.

O'Shaughnessy, B. (2009). The location of a perceived sound. In M. Nudds & C. O'Callaghan (Eds.), *Sounds & Perception. New Philosophical Essays* (pp. 111-125). Oxford: Oxford University Press.

Olofsson, J. K., & Gottfried, J. A. (2015). The muted sense: neurocognitive limitations of olfactory language. *Trends in Cognitive Sciences*, *19*, 314-321.

Pasnau, R. (1999). What is sound?. *The Philosophical Quarterly*, *49*(196), 309-324.

Pasnau, R. (2000). Sensible qualities: The case of sound. *Journal of the History of Philosophy*, *38*(1), 27-40.

Popova, Y. (2005). Image schemas and verbal synaesthesia. In B. Hampe (Ed.), *From Perception to Meaning: Image Schemas in Cognitive Linguistics* (pp. 395-420). Berlin: Mouton de Gruyter.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rauh, G, (2010). *Syntactic Categories: Their Identification and Description in Linguistic Theories*. Oxford: Oxford University Press.

Ronga, I. (2016). Taste synaesthesias: Linguistic features and neurophysiological bases. In E. Gola & F. Ervas (Eds.), *Metaphor and Communication* (pp. 47-60). Amsterdam: John Benjamins.

Ronga, I., Bazzanella, C., Rossi, F., & Iannetti, G. (2012). Linguistic synaesthesia, perceptual synaesthesia, and the interaction between multiple sensory modalities. *Pragmatics & Cognition*, *20*(1), 135-167.

San Roque, L., Kendrick, K. H., Norcliffe, E., Brown, P., Defina, R., Dingemanse, M., Dirksmeyer, T., Enfield, N., Floyd, S., Hammond, J., Rossi, G., Tufvesson, S., van Putten, S. & Majid, A. (2015). Vision verbs dominate in conversation across cultures, but the ranking of non-visual verbs varies. *Cognitive Linguistics*, *26*(1), 31-60.

Shen, Y. (1997). Cognitive constraints on poetic figures. *Cognitive Linguistics*, *8*, 33-71.

Shen, Y. (2002). Cognitive constraints on verbal creativity. In E. Semino & J. Culpeper (Eds.), *Cognitive Stylistics: Language and Cognition in Text Analysis* (pp. 211-230). Amsterdam: John Benjamins.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359-1366.

Spence, C., & Piqueras-Fiszman, B. (2014). *The Perfect Meal: The Multisensory Science of Food and Dining*. Chichester, UK: Wiley Blackwell.

Spivey, M. (2007). *The Continuity of Mind*. Oxford: Oxford University Press.

Strik Lievers, F. (2015). Synaesthesia: A corpus-based study of cross-modal directionality. *Functions of Language*, *22*(1), 69-95.

Strik Lievers, F. & Huang, C. R. (2016). A lexicon of perception for the identification of synaesthetic metaphors in corpora. *Proceedings of the 10th Language Resources and Evaluation Conference* (LREC), 2270-2277.

Stokes, D., & Biggs, S. (2015). The dominance of the visual. In D. Stokes, M. Matthen & S. Biggs (Eds.), *Perception and its Modalities* (pp. 350-378). Oxford: Oxford University Press.

Sweetser, E. (1990). *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge: Cambridge University Press.

Talmy, L. (2000). *Toward a Cognitive Semantics, vol. I: Concept Structuring Systems*. Cambridge: MIT Press.

Taylor, J. (2002). *Cognitive Grammar*. Oxford: Oxford University Press.

Tekiroğlu, S. S., Özbal, G., & Strapparava, C. (2015). Exploring sensorial features for metaphor identification. *Proceedings of the Third Workshop on Metaphor in NLP,* 31-39.

Tekiroğlu, S., Özbal, G. & Strapparava, C. (2014). Sensicon: An automatically constructed sensorial lexicon. *Proceedings of the EMNLP*, 1511-1521.

Ullmann, S. (1959 [1957]). *The Principles of Semantics* (2nd Edition). Glasgow: Jackson, Son & Co.

Urbanek, S. (2013). png: Read and write PNG images. R package version 0.1-7. https://CRAN.R-project.org/package=png

Vanhove, M. (2008). Semantic associations between sensory modalities, prehension and mental perceptions: A crosslinguistic perspective. In M. Vanhove (Ed.), *From Polysemy to Semantic Change. Towards a Typology of Lexical Semantic Associations* (pp. 341-370). Amsterdam: Benjamins.

Viberg, Å. (1983). The verbs of perception: a typological study. *Linguistics, 21*(1), 123-162.

Viberg, Å. (2001). The verbs of perception. In M. Haspelmath, E. König, W. Oesterreicher & W. Raible (Eds.), *Language Typology and Language Universals. An International Handbook*, 1294–1309. Berlin: Mouton de Gruyter.

Wickham, H. (2016). stringr: Simple, Consistent Wrappers for Common String Operations. R package version 1.1.0. https://CRAN.R-project.org/package=stringr

Wickham, H. (2017). tidyverse: Easily Install and Load 'Tidyverse' Packages. R package version 1.1.1. https://CRAN.R-project.org/package=tidyverse

Williams, J. M. (1976). Synaesthetic adjectives: a possible law of semantic change. *Language*, *52*, 461-479.

Winter, B. (2016a). Taste and smell words form an affectively loaded part of the English lexicon. *Language, Cognition and Neuroscience, 31*, 975-988.

Winter, B. (2016b). The Sensory Structure of the English Lexicon. PhD Thesis, University of California, Merced.

Winter, B., & Perlman, M., Perry, L. K., & Lupyan, G. (2017). Which words are most iconic? Iconicity in English sensory words. *Interaction Studies*, *18*, 433-454.

Wnuk, E., & Majid, A. (2014). Revisiting the limits of language: The odor lexicon of Maniq. *Cognition*, *131*, 125-138.

Yeshurun, Y., & Sobel, N. (2010). An odor is not worth a thousand words: from multidimensional odors to unidimensional odor objects. *Annual Review of Psychology*, *61*, 219-241.

Appendix A: Raw counts from the three word lists

		Sight	Sound	Touch	Taste	Smell	N
Norms	Adj	205	68	70	54	26	423
	Noun	336	42	14	6	2	400
	Verb	102	71	101	13	13	300
		643	181	185	73	41	
Author 1 (2015)	Adj	73	30	36	27	9	175
	Noun	49	107	8	14	15	193
	Verb	21	82	5	4	6	118
		143	219	49	45	30	
Sensicon	Adj	2074	1147	1121	1804	573	6719
	Noun	3005	2069	2186	2732	1091	11083
	Verb	800	761	764	961	452	3738
		5879	3977	4071	5497	2116	