■ *Simeon Floyd*
Universidad San Francisco de Quito
sfloyd1@usfq.edu.ec

■ *Lila San Roque*
Radboud University Nijmegen
Max Planck Institute for Psycholinguistics
l.sanroque@let.ru.nl

■ *Asifa Majid*
Radboud University Nijmegen
Max Planck Institute for Psycholinguistics
asifa.majid@let.ru.nl

Smell Is Coded in Grammar and Frequent in Discourse: Cha'palaa Olfactory Language in Cross-Linguistic Perspective

Abstract

It has long been claimed that there is no lexical field of smell, and that smell is of too little validity to be expressed in grammar. We demonstrate both claims are false. The Cha'palaa language (Ecuador) has at least 15 abstract smell terms, each of which is formed using a type of classifier previously thought not to exist. Moreover, using conversational corpora we show that Cha'palaa speakers also talk about smell more than Imbabura Quechua and English speakers. Together, this shows how language and social interaction may jointly reflect distinct cultural orientations towards sensory experience in general and olfaction in particular. [olfaction, sensory anthropology, Cha'palaa, Imbabura Quechua, English]

Introduction: Taking Account of Diversity in the Language of the Senses

The language of the senses is at the heart of longstanding debates about the degree to which language may influence or reflect differences in how people across cultures perceptually approach the world. The language of vision, and particularly of color, has provided a contentious example of how languages are claimed to develop in similar ways, based on universal cognitive and perceptual principals (Berlin and Kay 1969; Kay and Regier 2003; Kay, Berlin, and Maffi 2011). These generalizations have been problematized by Lucy (1997) for, among other things, not taking into account the specifics of individual languages, citing examples like the Hanunóo language of the Philippines in which "color" terms imply more than just color (as described in Conklin 1955; see also Wierzbicka 2005). Additional studies of minority languages like Yélî Dnye of Papua New Guinea (Levinson 2000) or Candoshi of Amazonian Peru (Surrallés 2016) have raised further questions about

the universality of claims about color terms in the face of cultural diversity (see Majid 2015a for further discussion).

In the domain of olfaction, universalist proposals are also beginning to be challenged by studies of diverse, lesser-known languages from around the world. Claims that smell is universally impossible to describe abstractly (Sperber 1975:115–16; Olofsson and Gottfried 2015; Yeshurun and Sobel 2010), and the related implication that olfactory language will be universally minimal across cultures (Lawless and Engen 1977; Stoddart 1990; Ackerman 1991; Wilson and Stevenson 2006; Olofsson and Gottfried 2015) are being confronted by research on the Aslian languages of the Malay Peninsula (Burenhult and Majid 2011; Tufvesson 2011; Wnuk and Majid 2014; Majid and Burenhult 2014; Majid and Kruspe 2018), among other languages with extensive abstract smell lexicons from northern Mexico (O'Meara and Majid 2016) to Africa (e.g., van Beek 1992; Blench and Longtau 1995; Hombert 1992; Storch and Vossen 2007), and Amazonian South America (Shepard 1999). These findings suggest—as has long been argued in sensory anthropology (e.g., Stoller 1989; Howes 1991; Classen 1997; Howes 2003)—that there is not a universal, biologically determined orientation to perceptual experience, but rather that sensory cultures are diverse (or feature different sensory "models"; Classen 1997).[1] This diversity suggests that the marginalization of smell may instead be a quirk of some Western cultures (Classen, Howes, and Synnott 1994; Majid 2015b; see also McGann 2017; on Western bias in selection of research subjects more generally, see Henrich, Heine, and Norenzayan 2010), related to historical transformations of the olfactory environments of Western urban contexts in connection to ideologies of hygiene and sanitation (see Corbin 1986 on France). People living outside such contexts may show greater sensitivity to smell, as one study comparing urban Germans to the Tsimane' of the Bolivian Amazon has shown (Sorokowska et al. 2013). Is the marginalization of smell actually a Western-specific cultural trait that can vary in different societies?

The role of language has not always been taken into account in studies of olfaction, even as the categories found in language provide a key entry point into how different cultural models approach sensory experience. Sensory anthropologists have critiqued how the "linguistic turn" in anthropology drew attention to the visual medium of the "text" at the expense other kinds of sensory experience (Howes 2005), but—as Majid and Levinson (2011) point out—focusing on indigenous linguistic expressions in diverse languages and cultures has precisely the opposite effect: it draws our attention to the diversity of cultural categories that partition the sensory domains in different ways (see also the call for attention to the linguistics of sensory language in Lucy 1997).

This study looks at diversity in olfactory language and culture through the example of the indigenous Chachi society of Ecuador, whose language Cha'palaa features specialized resources for talking about smell, with significant implications for debates of the supposed marginalization of olfaction. Cha'palaa bases its large set of olfactory terms on its grammaticalized "smell classifier", a type of classifier previously unattested in the typological literature (e.g., Aikhenvald 2003). We compare Cha'palaa's system of smell terms to that of its neighboring language Imbabura Quechua as well as to that of English, both to highlight what is distinctive about Cha'palaa, as well as to update previous accounts of the language of smell in these two other languages. We begin by looking at language structure, and then turn to language usage, drawing on natural speech corpora to examine specific cases of smell terms as used in social interaction and compare their relative discourse frequencies across the three languages. Finally, we discuss the potential significance of the linguistic facts about Cha'palaa for understanding diversity in cultures of olfaction across different societies.

Olfaction in Cha'palaa Speaking Society

Some preliminary observations based on long-term ethnographic study in Cha'palaa speaking communities indicate that smell may have special cultural significance in Chachi society. In a tropical rainforest environment where visual perception can be limited, olfaction is a source of information about plants, animals, human activity and weather conditions that local people are acutely aware of and pause to notice and comment about.[2] Olfaction is mentioned in traditional narratives such as tales of monsters that track humans by smell. References to smell are heard frequently in relation to cooking and hygiene products, and odors can play an important role in ritual, such as during the annual December *Niñu* ceremony; Floyd describes the experience:

> All of the surrounding communities had been gathered at the ceremonial house for days and nights of marimba and drumming that culminated in an event centrally involving perfume. Community members had explained to me ahead of time that perfume would be distributed to the crowd, but I had not realized that this meant that we would all be liberally soaked. Around midnight a group of people picked up the bottles that had been arranged on the altar, opened them, and poured perfume all over the upper body of each attendee, indicating the conclusion of the ceremonies. Fig 1.

The apparent significance of olfaction in Cha'palaa daily life and ceremony was further supported by the results of a 13-language comparison of sensory verbs in natural speech corpora. In this comparison, Cha'palaa ranked the sense of smell higher relative to all other languages except for Semai (San Roque et al. 2015), one of the Aslian languages noted for their olfactory language (Tufvesson 2011). In addition, developing a full account of Cha'palaa smell terms revealed that the most frequent terms *andyu* "fragrance" and *pudyu* "stench", are just two members of a larger set of abstract smell quality terms, all of which are formed with the morpheme -*dyu*, which grammatically classifies the resulting term as a type of smell.

This case of a smell classifier is significant not only because it adds to the cross-linguistic counterexamples to the statement that "there is no semantic field of smells" (Sperber 1975:116), but also because the typological literature on classifiers has yet to document classifiers based on smell. The different types of classifier systems seen in the world's languages are diverse in their grammatical expression (Corbett 1991; Senft 2000; Aikhenvald 2003), but semantically they are commonly based on core aspects like animacy, gender, shape, type of material, or usefulness (Seifart 2010:726–727), domains which are thought to be selected due to their high "cue validity" (Aikhenvald 2003:239–240; Seifart 2010:725–726), meaning a higher probability of

Figure 1.   Ritual communal use of perfume (*pindyupi*, 'sweetness-CL:SMELL-CL:LIQUID') in a Chachi community

inferring the properties of an object from a cue (Beach 1964; Reed 1972; Rosch et al. 1976). Seifart (2010:726) notes that cross-linguistically "properties that lack high cue validity, such as color, sound, feel, taste, or smell, do not play a role in any nominal classification system attested so far". Aikhenvald and Storch (2013) observe that smell-based grammatical categories are unattested in any linguistic domain; for example, they point out that there are no known olfaction-specific evidentials (2013: 3) and that "Neither hearing nor smell nor taste appear to be grammaticalised in other areas of linguistic structure: for instance, our cross-linguistic study of classifiers reveals these parameters never surface in any classifier type (Aikhenvald 2000), or in any other verbal or nominal category" (Aikhenvald and Storch 2013:36). While the Cha'palaa smell classifier plays a relatively minor role in the nominal lexicon compared to European or African gender markers, it represents the first known case of a nominal smell classifier, raising the question of whether such notable grammaticalization may be linked to the salience of smell in Chachi society.

A Cross-Linguistic Comparison of Cha'palaa with English and Imbabura Quechua

To better understand the special contribution of Cha'palaa we bring two further languages into the discussion: Imbabura Quechua and English. These languages are relevant for comparison with Cha'palaa in different ways: Imbabura Quechua because it is a neighboring language spoken in somewhat comparable conditions to Cha'palaa, and English because it is a well-known example of a Western language supposedly with few smell terms. In addition, there are a few updates from previous characterizations of both languages that are worth discussing in their own right.

In the case of Imbabura Quechua, it has been claimed that Quechuan languages have extensive smell language (Classen, Howes, and Synnott 1994), but it turns out that on closer examination there are minimal linguistic resources for talking about olfaction. Instead, what might look like dozens of distinct smell terms are actually complex words built up productively from extensive morphological processes, highlighting the importance of taking into account language typology in general and the agglutinative profile of the Quechuan languages in particular. English, on the other hand, is generally thought to have few linguistic resources dedicated to smell (Majid and Burenhult 2014:266), but here we make a more detailed survey of English smell quality terms and find the language to have a more extensive inventory than might be expected. First we address each language system in turn, and then turn to discourse data.

Linguistic Resources for Talking about Smell

There are two major ways that languages refer to olfactory experiences: words that refer to events of perceiving and emitting odors and words that refer to the qualities of odors. While each language's word class system is different, smell events are usually referred to by verbs, while smell quality words often fall into classes equivalent to nouns or adjectives.[3] Here we present the relevant facts about the linguistic resources for talking about smell in the languages under consideration, giving some brief background about each language, and then describing verbs for olfaction events, followed by smell quality terms.

The Language of Smell in Cha'palaa

Cha'palaa is one of the five Barbacoan languages of Ecuador and Colombia (Curnow and Liddicoat 1998), and is spoken by about 10,000 Chachi people in a tropical rainforest region near the Andes foothills in northwest Ecuador (Barrett 1925; Curnow and Liddicoat 1998; Floyd 2010, 2014a). Like many languages in the region, it is agglutinative and verb final, with large sets of verbal markers, including many evidentials. It stands out typologically in several ways, in particular due to its

tendency to rely on classifiers to transform its numerous unclassified roots into verbs or nouns. In the verbal domain, most predicates feature multiple roots in what is known as a "complex predicate" system (Dickinson 2002; Amberber, Baker, and Harvey 2010; Samvelian and Nash 2015). Such complex predicates can be observed in Cha'palaa's set of sensory distinctions (Viberg 1983). The verbs for "see," *kera*, and "hear," *mera*, occur as independent verbs, but the other senses occur together with *kera*, in the role of a general sensory verb.

In Table 1, Viberg's (1983) framework shows how, among Cha'palaa verbs, the "see" term acts both as an independent vision verb, and as a general sensory verb in combination with touch, taste and smell. Touch is literally "hold-grasp-*see*" and taste is "eat-*see*." Smell is the only one of the three with a dedicated root, *ish*, which combines with "see" or "grasp" for a more passive or active meanings (i.e., to perceive versus to actively sniff). While not a fully independent verb like "see" and "hear," the smell verb in Cha'palaa is thus a more specialized resource than the predicates associated with the touch and taste, which do not have dedicated roots for specific sensory meanings.

For talking about sensory qualities, Cha'palaa has several classes of words, some more noun-like and others more adjective-like, dedicated to referring to visual, tactile, and other sensory qualities. The terms for smell, and to some extent also taste, lean toward the nominal domain, in part because—like many nouns in the language—they are complex words derived from a root and a classifier. As discussed above, this is typologically relevant because classifiers typically single out categories such as gender or shape, but so far have not been attested to classify smells and tastes (Seifart 2010:726; Aikhenvald and Storch 2013:36).

The presence of classifiers in this specific domain is perhaps less surprising when one takes into account that classification is a basic part of many core areas of Cha'palaa grammar more generally. In Cha'palaa speech, many verbal and nominal constituents do not feature independent verbs or nouns, but rather combine sets of classifiers with abstract roots that usually cannot occur as main constituents on their own.[4] Smell terms are derived by an abstract root and the classifier -*dyu*—neither of which can occur independently—and are mainly used to attribute qualities of odors to things and situations through copulative predication (e.g., "It is stinky," "Flowers are fragrant") rather than noun modification (e.g., "fragrant flowers"). Other types of nominal classification in Cha'palaa involve shape classifiers for countable objects (spherical, oblong, elongated, etc.) that occur within nouns and agree with augmentatives and diminutives, numerals, and some adjectives.

While historically a distinction has been made between noun class systems (i.e., "gender") which involve agreement, and nominal classifiers, which involve individuation and derivation, Cha'palaa shows features of both, and Seifart and Payne (2007) note that in South American languages generally, classification systems often overlap between these categories. Since classification is a major way that independent words are formed in Cha'palaa, this may have provided a context for applying the

Table 1
Basic sensory verbs in Cha'palaa

	Activity	Experience	Source-based
SIGHT	*kera*		
HEARING	*mera*		
TOUCH	*ta'-ka'-kera* hold-grasp-see	no dedicated verb	
TASTE	*fi'-kera* eat-see	no dedicated verb	
SMELL	*ish-kera* smell-see		*ish-ka* smell-grasp

principals of classification in sensory domains not often included in classification systems cross-linguistically.

The Cha'palaa smell and taste classifiers combine with abstract roots, thereby creating independent sensory quality terms classified either as an odor (-*dyu*) or as a taste (-*mbu*). The most common combinations are with the positively and negatively valenced roots *a(n)-* and *pu-* (Table 2). While languages like English can create complex sensory terms (e.g., "good smell" or "bad smell"), the component referring to smell and the component referring to quality are independent words. In Cha'palaa, neither component is independent; the "good/bad" element and the "smell" element grammatically depend on each other and must occur together.

The sensory classifier for smell is more productive than that for taste, and in addition to the two positive and negative valenced terms, it combines with other roots to create additional smell quality terms (described below), while further taste terms have not yet been attested. *Pudyu*, "stinky" and *andyu*, "fragrant" are the most frequently heard smell terms, aligning with claims that a basic pleasantness/ unpleasantness distinction structures olfactory experience (Yeshurun and Sobel 2010). In addition, speakers of Cha'palaa make further more specific distinctions using abstract smell quality terms, pinpointing other dimensions of odors. Table 3 lists all currently identified terms and their approximate meanings.

The set of roots classifiable as smell qualities (*pu-, an-, pin-*. . . etc.) is "closed" in the sense that it was impossible to create new combinations that speakers would accept (e.g., combining *kucha*, "dog," with -*dyu* to say "dog smell," and so on). They come from a large class of grammatically unclassified roots that are not inherently verbs, nouns, or adjectives, but that are derived as such through combination with classifiers. Some of the same roots occur in morphologically complex terms not referring directly to smell—for example, *pin* is used in a number of distinct terms for "sweet" things; *lush* may be related to the color term "grue"; and *juluj* occurs in a term for "dandruff"—but as stated before the roots cannot occur in an underived state, so it is unclear whether a unified semantic meaning for such roots can be uncovered. Although part of a closed set, new instances of smell terms continue to turn up in elicitation and natural speech data, suggesting that Table 3 may not be fully exhaustive. The present number of 15 attested terms in Cha'palaa is relatively large in cross-linguistic perspective, and contrasts with neighboring Imbabura Quechua's lack of smell quality terms.

The Language of Smell in Imbabura Quechua

The Quechuan languages originated in Peru and spread throughout the Andes over many centuries, especially in connection with the expansion of the Inca Empire between the 14th and 16th centuries (Torero 1964; Cerrón-Palomino 1987; Adelaar and Muysken 2004). In the territory that would become Ecuador, local societies were strongly influenced by Inca language and culture in this period (Salomon 2007). In the Ecuadorian Andes today there are many different local ethnic groups who speak regional varieties of Ecuadorian Highland Quechua, a distinct language from other Quechuan varieties. The specific dialect that we examine here is known as Imbabura Quechua (Cole 1985), spoken in the northern Ecuadorian Andes in villages and rural

Table 2
Cha'palaa sensory classifiers

valence	classified as smell	classified as taste
positive	*an-dyu* "fragrant"	*a-mbu* "delicious"
negative	*pu-dyu* "stinky"	*pu-mbu* "distasteful/disgusting"

smell term	Meaning
pudyu	stench, generally negative; also used in compounds like *fi'pudyu* (mouth stench), *pijpudyu* (mud stench), *tupudyu* (dirt stench), *chijpudyu* (*chijdyu* plant stench), etc.
andyu	fragrance; good food, generally positive
pindyu	sweet smell; perfume
pe'dyu	rotten smell; rotting meat, excrement
sendyu	fishy or metallic smell; raw fish or meat, blood
jedyu	strong, acrid smell; gasoline
jemeedyu	similar to *jedyu*, but specifically used for alcohol
lushdyu	smell of fresh cut plants or plant seeds
julujdyu	smell of dry and dusty air; e.g., accompanying sources such as dry wood, or other objects dusted with dry earth or plant matter
niwishdyu	smell of burning, smoke
lajdyu	pleasant smell of some foods; e.g., squash, fruit, baking bread
wishdyu	smell of urine
pijdyu	smell of some wild animals; e.g., wild boar
chijdyu	smell of burning hair
mira'dyu	slightly rotten, rancid smell; food going bad

settlements in the valleys and mountain slopes of the province of Imbabura. Locals of this area engage in different types of agriculture at different altitudes, and participate in a long tradition of weaving and handcrafts. It is a cooler and drier environment than the lowland forests where the Chachi people live, and closer to urban centers, but historical connections of trade and migration have led to considerable shared cultural elements between the Chachi and the speakers of Imbabura Quechua (Floyd 2009, 2010). Locals use incense and distinct smelling plants for some ritual purposes; there is no obvious reason to assume that smell be particularly minimized or particularly salient relative to Cha'palaa speaking society, one way or the other.

Typologically, Imbabura Quechua is similar to Cha'palaa in that it is an agglutinative, verb-final language with extensive verbal morphology, but notably, it differs from Cha'palaa in that it does not employ classifier systems, neither in the verbal nor nominal domain; the language does not have any grammaticalized classifiers at all, meaning it cannot rely on them for expressing sensory experience as in Cha'palaa, and must use other strategies.

With respect to their verbs for olfaction, the Quechuan languages are worth revisiting because they have been previously described as featuring extensive sensory language, with respect to olfaction in particular (Classen, Howes, and Synnott 1994: 112). Using a 17th-century dictionary of a Peruvian Quechua language in the same major Quechuan subgroup as Imbabura Quechua, Classen et al. (1994:112) identified a number of terms with translations that appear to involve very specific types of smelling, concluding that "the extent to which the Incas took care to distinguish between kinds of smelling indicates something of the attention that they paid to the olfactory process and its different effects." However, while dictionary entries appear with very specific translations like "for a group to smell something together," "to secretly sniff out what is being planned," "to make someone smell something," the colonial-era Spanish priests who compiled such documents often did not distinguish verb roots and fully conjugated verbs that are essentially whole sentences. In these examples the complex meanings do not come from the verb roots *azna* "to smell," *mutqqui* "to smell/sniff," or *sama* "to breathe," but rather from the additional morphemes added from the Quechuan languages' large inventory of verbal markers.

For example, *mutqquichini*, "to make someone smell something" is simply a combination of the root "to smell (actively)" and a causative marker *-chi* which could be applied to any verb, and which is not a dedicated resource for talking about smell. This means that the actual unique mono-morphemic smell terms are considerably fewer than the total dictionary entries that mention olfaction in their definitions.

While there is some variation among Quechuan languages, they appear to mostly share cognate forms of the same smell verbs, and these do not appear to number more than two or three. The languages do not lack expressive capabilities, given the general morphology discussed above, but this is not the same thing as having specialized language for talking about olfaction. By that standard, Quechuan languages actually appear undistinguished in their smell lexicon. Table 4 reviews the basic sensory verbs in Imbabura Quechua.

Imbabura Quechua does not have distinct verbs for activities versus experiences for vision or hearing (e.g., *look* versus *see*; *listen* versus *hear*), unlike English. Further distinctions in meaning can be achieved morphologically, such as by adding a reflexive marker *-ri* for some source-based sensory events. Taste is particularly sparse in terms of resources, with only a verb *mali* for "to try food/drink"; otherwise simple phrases like "it is good/bad (to eat/drink)" may be used to refer to taste. There is a dedicated verb for smell that can be used for both active and source-based smell events, *ashna*, as well as another verb *mutiji* that can be used to refer to the action of "sniffing" to perceive a smell.

Specialized vocabulary for talking about smell qualities is also limited in many Quechuan languages, and in Imbabura Quechua it does not exist at all. Speakers either use source-based descriptors ("garlic smelling") or use more general or multi-sense terms ("bad (smelling)," "sweet (smelling)"). Since there are no smell quality terms, the two smell verbs make up the entire smell lexicon of Imbabura Quechua.

The Language of Smell in English

English is a Germanic language with myriad dialects (Baugh and Cable 2012), and has native speakers numbering between 300 and 400 million, with many more non-native users in different settings around the world (Crystal 2003). Here, however, we consider the English of largely urban speakers of UK and US English (see descriptions of corpora below).[5] Typologically distinct from highly agglutinative languages like Cha'palaa and Imbabura Quechua, English is largely analytic with minimal inflection (Quirk 1980). The sociocultural environments of English are varied far beyond the scale of the small-scale communities where Cha'palaa and Imbabura Quechua are spoken.

Based on the analysis of Viberg (1983), the resources of US/UK English for talking about sensory experience are made up of—in the verbal domain—at least one dedicated verb for each of the senses (Table 5). Vision and audition feature verbs for activities (*look [at]*, *listen [to]*) and experiences (*see*, *hear*), as well as roots for "copulative" constructions that take the percept as the subject (*look* and *sound*). Taste and smell have one major verb each, while touch is expressed with either *touch* or

Table 4
Basic Imbabura Quechua sensory verbs; some verbs take the REFLEXIVE *-ri*

	ACTIVITY	EXPERIENCE	SOURCE-BASED
SIGHT	*riku*		*riku-ri*
HEARING	*uya*		*uya-ri*
TOUCH	*taka-ri*		
TASTE	*mali*	no dedicated verb	
SMELL	*mutiji*	*ashna*	

Table 5
Basic sensory verbs in English, adapted from Viberg (1983:128)

	ACTIVITY	EXPERIENCE	SOURCE-BASED
SIGHT	*look at*	*see*	*look*
HEARING	*listen to*	*hear*	*sound*
TOUCH	[*touch*]	*feel*	
TASTE	*taste*		
SMELL	*smell*		

feel.[6] A feature of the English vocabulary—distinct from Cha'palaa and Quechua—is that many of the core terms (*look, sound, touch, feel, smell, taste*) can occur as nominal as well as verbal roots, without any overt derivation (e.g., the same form occurs in the nominal phrase *a suspicious* **look** and the verbal phrase *you* **look** *suspicious*).

As for smell specific quality terms in English, these have generally been understood to be limited to the negatively valenced terms *stench/stinky* and the positively valenced terms *fragrance/fragrant*. Speakers of English often use source-based terms like *rotten* or *fishy* to describe many smells. However, there are a few more terms in English that might be considered dedicated smell terms (e.g., *musty*), and we consider a broader set of terms in our discourse frequency comparison below; English appears to be a bit more expressive in this area than is generally thought.

Olfactory Language in Discourse

What does it mean for Chachi society that Cha'palaa has special linguistic resources for referring to smell? The cultural implications of elaborated lexicons in specific domains has been long debated (compare Pullum 1991; Martin 1986; Regier, Carstensen, and Kemp 2016). However, it is unclear what evidence could show that embellished language in specific domains implies cultural salience. One cross-linguistic approach for studying social behavior independently from language structure is to turn to language usage. A parallel can be found for vision. In a discussion of color terms, Goodwin (1997) pointed out that the comparison of the language of vision across cultures is based primarily on the lexical systems for referring to color categories, whereas the complementary study of the usage of color terms in social interaction can provide important parallel insights.[7] Natural speech recordings represent concrete examples of how people respond to, and socially negotiate, sensory experience in daily life; and offer a more ecologically valid method to complement controlled elicitation. With large enough corpora, it is also possible to gain insights about the relative hierarchy of the senses by comparing discourse frequency across languages (e.g., San Roque et al. 2015; Winter 2016). In this section we look beyond language structure, and compare the frequencies and types of expressions of olfactory experience found in discourse.

Minority Language Corpora

Transcribed video corpora for largely unwritten minority languages like Cha'palaa and Imbabura Quechua are only recently becoming available, making it possible to compare across languages, including those such as English for which large corpora already exist. This comparative approach adds a cross-linguistic dimension, and highlights data from minority languages that have been under-represented in corpus linguistics. At the same time, including data from these languages also presents some challenges: "Large" corpora for minority, unwritten languages will be much smaller than "large" corpora for languages with speakers in the millions like English. While for English we could access a range of publicly available corpora representing

hundreds of hours of transcribed recordings, for the two indigenous Ecuadorian languages the recordings are the result of primary fieldwork by Floyd and a small team of collaborators, who must follow painstaking procedures of transcription, with minimal technical support and few existing descriptive language resources, in order to reach several dozen hours.[8] While much smaller than English-language resources, these corpora are nevertheless much larger than what is generally available for minority unwritten languages.

Information on Corpora

The Cha'palaa and Imbabura Quechua corpora were collected under similar conditions, with Floyd collecting video in primarily informal village and household contexts, and transcribing and translating the recordings with the assistance of native speakers in the ELAN software program (Sloetjes and Seibert 2016). The corpora are multipurpose, and were collected according to standard methods for long-term ethnographic and collaborative documentation of endangered languages (e.g., Gippert, Himmelmann, and Mosel 2006), as well as for general research use. The still-growing Imbabura Quechua corpus was approximately 150,000 words when consulted, while the Cha'palaa corpus was more than three times as large, with about 500,000 words.[9] During fieldwork the camera was often left unattended to capture spontaneous conversation, while in some cases the researcher remained present to record semistructured narratives or ethnographic interviews. All of the recordings can be considered informal and noninstitutional conversation, which is arguably the most cross-linguistically comparable linguistic data (Stivers et al. 2009; Dingemanse and Floyd 2014). The Imbabura Quechua corpus was recorded in the contexts of villages set in agricultural land among lakes and high mountains, adjacent to several provincial cities. The Cha'palaa corpus was recorded in riverside villages set in humid tropical rainforests remote from any urban areas. Figs 2,3.

In this study we included three different sources of spoken English, each with their advantages and disadvantages. First, the Spoken section of the Corpus of Contemporary American English (COCA, Davies 2008–) provides a very large database of unscripted spoken English that is auto-tagged for part of speech, numbering approximately 109 million words. This is a tremendous resource and the volume of data is impressive. However, the corpus is largely composed of television and radio talk shows and interviews, and thus might not represent typical, informal face-to-face interaction. The second source of data was the freely available Santa Barbara Corpus of Spoken American English (Du Bois et al. 2000–2005), an audio-recorded and fully transcribed corpus of approximately 22 hours of English conversation (approximately

Figure 2.   Screenshot from Imbabura Quechua corpus of people engaging in conversation and informal household tasks (shucking corn) in an Andean village.

Figure 3.   Screenshot from Cha'palaa corpus of people engaging in conversation and informal household tasks (basket weaving) in a village in the lowland rainforest.

300,000 words). In general, this material fits well with the specifications of the Cha'palaa and Quechua material, but also includes a small amount of data that is not so comparable, such as telephone (as opposed to face-to-face) conversations, and institutionalized interactions (e.g., a college lecture).

The third source was a selection of about 60,000 words from the Rossi Corpus of English (Rossi and Kendrick 2013) comprising video recordings made in informal settings in England (Figure 4). This is a better match to the Cha'palaa and Imbabura Quechua corpora methodologically, but is smaller than the other corpora. Figures 2–4 illustrate the types of informal contexts represented in the corpora.

We made two different frequency comparisons based on these corpora: the first of the relative frequencies of sensory verbs, and the second of quality terms for olfaction.[10]

Comparison of Sensory Verbs

For all three languages, we counted all occurrences of the basic verbs (see Tables 2, 4, and 5) for the five senses.[11] Overall, our analysis of verb frequency recapitulates the findings of San Roque et al. (2015): vision verbs dominate in conversation, while touch, taste, and smell are less frequent than would be expected if all senses were

Figure 4.   Screenshot from Rossi Corpus with speakers of UK English engaging in informal conversation in an urban setting.

equally salient. The trend is further confirmed by the inclusion of Imbabura Quechua in the comparison, which was not included in San Roque et al. (2015). We also confirmed the unusual rank order of Cha'palaa—i.e., smell has a much higher ranking than in other languages—using a substantially larger corpus. As the sample of Cha'palaa included in this study was almost 100 times larger than in San Roque et al. (2015) (95 hours compared to 1 hour), this ranking appears to hold for Cha'palaa across recordings and speakers.[12] For Imbabura Quechua and English, smell was ranked second lowest and lowest, respectively.

When speakers of Cha'palaa, Imbabura Quechua and English did refer to smells in conversation there were some interesting differences, too. In Cha'palaa the verb "smell" was used primarily to talk about smell events, since references to abstract qualities use the classifier terms. In **Example 1** a girl mentions that a specific container of water was not fit to drink, and her mother asks her if she had sensed this by smelling it.

Example 1: Cha'palaa CHSF2012_08_03S1_922300

G: *pi   juntsa-na-a   man-ka-'   kush-nu   ke-daa*
   water   that-ACC-FOC   again-get-SR   drink-INF   do-HORT
   'Let's all drink *that* water (and not the dirty water).'

M: *ish-ka-'   kush-yu   ñu-ba*
   smell-get-SR   drink-EGO   2-also
   'Did you also smell (the dirty water) when you drank it?'

Because there are no abstract smell quality terms in Imbabura Quechua, the smell verb was used not only to talk about events, but often occurred in combination with nouns for smell sources to talk about qualities. **Example 2** shows a reference to smell from the Quechua corpus linked to a specific nominal source: a man and his wife discuss how a healer cured his chronic nose bleeds by boiling cow hooves in water and making him smell the vapors from the tea.

Example 2: Imbabura Quechua QUSF2015_12_02S1_2032594

S: *wakra tzillu   chay chay ashna-ju   chay chay   zumo-gu*
   cow   nail/hoof   that that smell-CONT   that that   liquid/juice-DIM
   'Cow hoof, that, that, smelling, that that herbal tea,'

M: *kusa-ri-shka*
   cook.over.fire-REFL-PTCP
   'Cooked over a fire.'

S: *ari   chay kusa-ri-shka-gu-ta-ka*
   Yes   that cook.over.fire-REFL-PTCP-DIM-ACC-FOC
   'Yes, cooking that over a fire...'

In English, the verb *smell* was commonly used both for reference to olfactory experiences, as well as in combination with source terms. Speakers often evaluated smells as good or bad, as in **Example 3** in which two friends are baking a cake together and one comments on the smell when she takes it out of the oven.

Example 3: English RCE26aCatan_3245994

L: ((takes cake out of oven))   *it   smells   good*

In English the verb *smell* is sometimes also used metaphorically, for example, in phrases such as *She seems to smell victory* (COCA/SPOK/ABC). This type of usage was not prominent in the Cha'palaa or Quechua corpora, as we discuss further below.

Differences in the use of smell verbs across these three languages may be conditioned in part by the broader structure of the olfactory lexicon. This is seen most clearly in the comparison of Cha'palaa and Quechua. As Cha'palaa has an elaborate set of terms for smells, the predicates with the different smell verb options are mainly used to talk about acts of smelling, but in Quechua the lack of dedicated smell terminology terms means the basic smell verbs are used for both acts of smelling and specific smell qualities.

Comparison of Smell Quality Words

We wished to know whether Cha'palaa speakers use their specific smell terms more often than English or Quechua speakers use theirs. While not differing greatly with respect to the number of their basic smell verbs, the number of olfactory quality terms in each of the three languages varies greatly. Cha'palaa has 15 abstract terms using the classifier *-dyu* (see Table 3), while Imbabura Quechua has no dedicated smell quality terms. For English we compiled a comparable list of smell terms (collated over many years through dictionary searches, literature searches, etc.), and focused on those terms that convey a specific smell quality.[13] The final list was based on the root terms: *acrid, aroma, fragrant, musty, noisome, odor,*[14] *pong, pungent, rank, reek, smelly, stench,* and *stink.*

Since Quechua has no smell quality terms, Cha'palaa smell quality terms are obviously more frequent, but trivially so. However, we also found that Cha'palaa speakers referred to smell qualities more often than English speakers.[15] Interestingly, Cha'palaa speakers were equally likely to mention unpleasant *pudyu* (39% of all smell quality tokens) or pleasant *andyu* (33%) and *pindyu* (10%) smells; whereas English speakers' talk was dominated by unpleasant smells. *Stink,* or some variant thereof, accounted for 53% of all smell quality words, followed closely by *smelly, stench, reek* and *rank,* which together accounted for 82% of tokens. In contrast, *fragrant* (4%) and *aroma* (4%) rarely appeared.

Overall, the contexts in which smell quality terms were found in the corpora were diverse. One context for the Cha'palaa terms can be found in the description of intense smells in the environment. In **Example 4**, the grandmother (GM) notices that her granddaughter (GD1) is reacting to a strong smell in the vicinity, and comments to her daughter-in-law that one of the crabs bought in the market that morning had died in transport in the canoe, and was now rotting.

Example 4: Cha'palaa CHSF2011_01_10S1_1251291

GD1:　((child reacts to smell and rubs nose))
GM:　*Pe'-dyu-n-ke-e-ti-shu,*　　　　　*main　tute-n-shu　kaa　suya　　　nelu*
　　　rotten-CL:SMELL-PFV-do-do-say-IRR　one　kill-PFV-IRR　DIM　daughter.in.law　crab
　　　'She must be saying its making a rotten smell, a crab was killed, little daughter-in-law.'

In the more abstract discussion shown in **Example 5**, two men explain that the odor of wild boars is perceived as unpleasant (that is, *pudyu*) by people, but fragrant (*pindyu* or *andyu*) by boars, expressing what might be seen as a kind of olfactory "perspectivism" (Viveiros de Castro 2011).

Example 5: Cha'palaa CHSF2015_01_31S3_938952

Y　*awen-dyushu　pu-dyu　　　　ki-shu-juntsa-a　　　ya-che-e*
　　great-boar　　stench-CL:SMELL　do-IRR-that.one-FOC　3-DAT-FOC
　　'For the large boar, that stench, for him...'

E   *pin-dyu-pi*
     sweet-CL:SMELL-CL:LIQUID
     '(It's like) perfume.'

Y   *jee   pin-dyu-pi-ee   de-ti   (chinba)   ti-ee-ti,   an-dya-ya   ju*
     Yes   sweet-CL:SMELL-CL:LIQUID-FOC   PL-say   (unclear)   say-DCL-say   fragrant-CL:SMELL-RED   be
     'Yeah (like perfume) they say, they say it is fragrant (for the boar).'

E   *serio*
     'Seriously.'

Y   *pu-dyu-jtu*
     stench-CL:SMELL-NEG
     'It does not stink.'

The English smell quality terms in the COCA Spoken corpus tended to recur in relation to certain stimuli and situations. For example, *acrid* was typically used with descriptions of burning and smoke, often occurring to set a scene of hazard and toxicity in relation to dramatic events. In **Example 6**, a survivor from a plane crash describes the situation inside the cabin.

Example 6: English COCA/2000/SPOK/NBC_Today

A: *The   fumes   were   acrid,   and   you   just   felt   like   you   were   being   poisoned.*

By contrast, a dominant use of *aromatic* is in reference to food and drink; in the COCA corpus this word is typically found in morning "lifestyle" segments of weekend talk shows during product discussions and/or cooking demonstrations. More rarely, *aromatic* turns up in science and news programming in relation to chemical compounds.

It is striking how often English smell terms are used with a metaphorical meaning rather than a literal olfaction meaning. For example, while the word *reek* is often used concretely to describe negatively evaluated odors (especially those concerned with the body, such as sweat, vomit, or decay) around one third of the occurrences of *reek* in this corpus were figurative, as in examples such as *Prisons, of course, reek of idleness* (COCA/1998/SPOK/NPR_Sunday) or *it reeks of political favoritism* (COCA/2009/SPOK/CNN_Grace). Even *stink*—the most frequent smell quality term in the Spoken COCA corpus—shows a high rate of non-literal use (e.g., as in, *I will not make a stink over it* [COCA/1993/SPOK/ABC_Brinkley]; or *They are stinking fakes* [COCA/2005/SPOK/CNN_Crossfire])

Figurative uses of smell terms do not make a showing in the Cha'palaa corpus, suggesting a subtle difference in the common uses of smell quality terms across the English and Cha'palaa data. While Cha'palaa speakers are able to select from a detailed inventory of terms to describe actual, experienced odors, the COCA corpus shows English speakers using a few terms to figuratively describe situations and actions as reprehensible or delightful.

Conclusion

Returning to the question of the potential sociocultural significance of the particularly complex language of smell in Cha'palaa, we can now say a few things about how linguistic resources and language usage may be connected. Looking beyond the grammatical properties of the system to its use in discourse uncovers results for frequency that parallel those seen in the comparison of linguistic resources: Cha'palaa stands out from the other two languages with respect to the language of smell by both standards. Speakers of Cha'palaa have more complex linguistic resources

dedicated to talking about olfaction than speakers of Imbabura Quechua and English, and they also talk about smell more often than speakers of those languages.

Does this mean that usage is shaping grammar or that grammar is structuring usage? Does having many smell terms imply a cultural emphasis on olfaction among the Chachi? There are long-standing debates over whether usage frequency is the cause or the result of grammaticalization (Hopper 1996; Hopper and Traugott 2003; Bybee 2007). While the causal directionality is unclear here, the number, and degree of conventionalization, of smell terms does correlate with frequency of use in the languages considered. A cultural preoccupation with smell in Chachi society may motivate the development of smell terminology, or acquiring the many smell terms in the language may lead people to speak relatively more frequently about olfaction.

A more holistic way of thinking about it is that language and social interaction mutually "boost" each other in directions that may vary across sociocultural contexts, and in Chachi society notably include an orientation towards olfaction. This suggests a loosely relativistic account of language and sensory orientations, at least in terms of speech production and comprehension ("thinking-for-speaking"; Slobin 1987).[16] The extent to which there are further implications for language-culture relationships depends on whether discourse frequency can be taken as a real proxy for cultural prominence, which is difficult to determine (see also San Roque et al. 2015). But the type of relativism suggested by the correlation between grammatical complexity and discourse frequency is not a simplistic scenario (e.g., "many words for snow"), but rather one in which language and habitual social practices are deeply interrelated and culturally transmitted together.[17] This scenario stands in contrast to the claims of universal marginalization of smell discussed above, as what we observe is considerable cultural diversity and variability.

We do not find evidence that the language of perception is infinitely variable however; both Viberg's (1983) structural hierarchies as well as rankings of discourse frequencies of sensory verbs shown by San Roque et al. (2015) are replicated here. These results confirm that vision tends to be prominent in sensory language cross-linguistically. In a response to Howes and Pink's (2010) critique that Ingold's (2000) discussion of sensory perception focuses too much on vision at the expense of the other senses, Ingold replied with the observation that "eyesight is quite obviously important to most human beings everywhere" (Ingold 2011:316). We have focused on smell here, but an in-depth look at Cha'palaa visual language would find a highly elaborate color term system (Floyd 2014b:97–103) rivaling anything seen in a "visual-centric" Western language. However, this does not exclude Cha'palaa from having complex olfactory language as well (see also Majid and Burenhult 2014; Majid and Kruspe 2018). Vision may be universally prominent, but that does not mean that cultural orientations toward perception cannot be diverse.

Since both universalistic and culturally relative forces are at play in shaping the language of perception, better than asking whether or not smell is universally marginalized is to ask to what degree species-general and culturally variable factors complementarily account for what we see across languages and cultures. Natural speech corpora for unwritten minority languages like the ones analyzed here have an important role to play in seeking answers to these questions by helping to ground our theories in what people across communities actually say and do.

Notes

English corpora, we thank their curators for making them publicly available, and especially Mark Davies for further advice concerning COCA. We also gratefully acknowledge comments on drafts of this paper by Ewelina Wnuk, Georgia Ennis, and Chip Zuckerman. We also especially thank members different communities in Ecuador and the UK for generously sharing their time and thereby contributing the valuable data analyzed here.

1. Research in linguistic anthropology, on the other hand, while touching on sensory perception in different ways, has seldom approached the linguistic forms associated with sensory perception systematically as a way into different sensory cultures. Feld's detailed ethnographic work on Kaluli approaches to sound and the sense of hearing (1979) is one of the few studies that explicitly deals with language and communicative expression in relation to a specific sensory domain. Notably it does this by spanning ethnomusicology and linguistic anthropology. Other relevant work can be found in *The Senses and Society* (e.g., *The Senses in Language and Culture* [Majid and Levinson. 2011] and a recent volume on olfaction (Barkat-Defradas and Motte-Florac 2016).

2. Majid and Kruspe (2018) found that two closely related communities located in the tropical rainforests of Malaysia, nevertheless, behave differently with respect to their use of olfactory language. This shows that environmental determinism is unlikely to be at play here. Still, the affordances of specific sensory environments where smell may be salient and vision limited are worth considering in the larger typological context.

3. Languages are diverse and smell quality terms are not universally nouns or adjectives; for example, in Huehuetla Tepehua they are ideophones (Kung 2007; see also O'Meara, Kung, and Majid forthcoming), while in some African languages it has been claimed they are a distinct class (e.g., Blench and Longtau 1995; Storch 2013).

4. On verb classification see Floyd (2014b, forthcoming); see also Dickinson 2002 for a description of a similar system from a related language that influences this analysis.

5. English is spoken in very diverse settings but over 80% of the populations of major English-speaking countries like the United States, UK, Canada, Australia, etc., live in urban areas. Source: http://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS?year_high_desc= true

6. Viberg (1983) does not include *touch* as a core perception verb, presumably because it can be argued to be a verb of contact rather than perception, since inanimate objects can be described as touching each other (e.g., *the cup touched the saucer*), but the experience of tactile perception is not involved (see, e.g., Johansen 1997:178–179). However, lexicographic sources define the verb *touch* as having perception as part of its primary meaning (see OED, Merriam Webster), and it is often used in this sense, suggesting that *touch* could be included as a core perception term. We follow the more inclusive path and include it in our study for consistency (e.g., the verb *see* also has several nonperceptual meanings, but this does not exclude it from being considered a core perception term). This choice also allows for comparison with languages that do not have a *feel*-like term, but typically express touch using predicates that also have "contact" semantics.

7. Several transcribed cases of references to smells in the three languages are discussed in here (see examples 1–6); a more detailed microanalysis of further cases along the lines of Goodwin (1997) would be an important future study to complement this one.

8. Most data studied in corpus linguistics is currently from languages that have extensive written resources (McEnery and Hardie 2011:11–12), and only a few minority languages have extensive sources available in online digital text (see Scannell 2007; Sene-Mongaba 2015; Llitjós, Aranovich, and Levin 2005), which in turn makes it more difficult to develop tools for facilitating data processing (McEnery, Baker, and Burnard 2000). The goals of corpus linguistics and those of the language documentation projects for which most minority language corpora are created may not align (see Cox 2011), and the types of annotation prioritized for one may not always be applicable to the other. Full glossing and part-of-speech tagging is often too much to hope for; the main method we use here is exporting lists of specific strings of text, and manually editing for any errors on a case-by-case basis.

9. Because agglutinative South American languages have a higher morpheme-per-word ratio than English, we estimated corpus size by both by word count and character count, and found the same results by both measures.

10. For Cha'palaa and Imbabura Quechua, we used the ELAN software to extract word lists and frequencies from the corpora. For the English corpora, the COCA Spoken data was searched online using the BYU corpus portal http://corpus.byu.edu/. For the figures reported in this paper, searches were conducted using individual words (i.e., as opposed to lemmas, customized word lists, or published frequency figures) during the period May–July 2016 (verb

frequencies) and June 2017 (quality terms). For the Santa Barbara Corpus, frequency counts and hand-checking were done using CLAN and Terminal, and for the Rossi Corpus, frequency counts and hand-checking were done using ELAN and Terminal. Individual words were found to be more reliable and transparent search strings than lemmas or lists. For example: the COCA corpus does not automatically include the string *smelt* as an instantiation of the lemma SMELL, but *smelt* does occur with smell-related meanings in the corpus. BYU corpus managers further advised that the (typically slightly lower) results obtained through single-word searches were likely to be more conservative than list searches, as tokens that have been identified as questionable or potentially mistagged are not returned. However, the differences between the two methods are very small.

11. Since some English terms (e.g., *sound*) can occur as either nominal or verbal roots, in contrast to Cha'palaa and Quechua, we took the following steps to exclude nominal roots from the English data, and thus make a more appropriate comparison. Firstly, any item marked with a suffix that is usually only applied to verbal roots (that is, *-ing*, *-ed*, *-er*) was counted as an occurrence of the verb root. However, without context, unmarked roots (e.g., *look*) and roots marked with a suffix *-s* (e.g., *tastes*) are ambiguous between nouns and verbs. Certain forms are also ambiguous with other, nonverb meanings (e.g., *felt* as a kind of cloth). To deal with these problems we identified which English terms were highly polysemous (i.e., likely to occur as nonverbs) by checking the occurrence of potentially ambiguous forms in the COCA (Spoken) corpus, which is automatically tagged for part of speech. We regarded an item as highly polysemous if it was tagged as a verb less than 95% of the time, giving 10 terms (*look, sound, sounds, touch, touches, smell, smells, smelt, taste, tastes*). We identified which occurrences of these 10 items in the corpora were indeed used as bare or inflected nominal roots rather than as verb roots, using automatic tagging in the COCA corpus and hand-checking in the Santa Barbara and Rossi corpora. The nominal occurrences were then excluded from the final counts.

12. We began the verb comparison comparing the rank order of sensory verbs based on their frequency. We found that all the English corpora—the COCA, Santa Barbara and Rossi corpora—correlated highly with each other at $r = 1.0$ ($N = 5$) using Kendall's tau-b correlations. That is, across all English corpora, vision verbs were used most often followed by audition, touch, smell and taste verbs. There was also a correlation between the English and Imbabura Quechua corpora $r = .80$, $p < .05$. This correlation was smaller than among English corpora because the frequency of taste and smell were reversed in Quechua with taste appearing slightly more frequently. However, the rank order of the senses in Cha'palaa did not correlate with English $r = .60$, $p = .14$ or with Quechua $r = .40$, $p = .33$. In Cha'palaa, the rank order of the senses was: vision, audition, smell, taste, and then touch: smell was ranked more highly than in the other languages. This effect appears to be driven both by the high rate of smell verbs and by the vanishingly low frequency of touch verbs relative to smell and taste. Alternatively, instead of comparing the rank order of the senses, we could ask whether people are equally likely to talk about each of the senses with verbs in conversation. If each sense was as salient as every other, they would populate conversations equally as often. That is not the case for any language: English: COCA $\chi^2(4) = 908,046$, $p < .0001$; Santa Barbara $\chi^2(4) = 3,865$, $p < .0001$; Rossi $\chi^2(4) = 748$, $p < .0001$; Quechua $\chi^2(4) = 1,223$, $p < .0001$; Chapalaa $\chi^2(4) = 11,649$, $p < .0001$. In all the languages, basic vision predicates appear far more often than would appear if the senses were equally frequent, with touch, taste and smell appearing far less frequently than would be expected by chance.

13. This excluded generic terms such as *smell* (noun) and *odor* (noun), as well as terms such as *whiff* (which appears to indicate the brief duration of the thing smelled, rather than a specific odor quality). However, terms such *smelly* (adj) and *odorous* (adj) were included as they appear to indicate a negative odor quality. The three authors (all native speakers of different English varieties) made judgments on all cases and only those terms with majority consensus on a meaning indicating smell quality were included in the final list. We aimed to be as inclusive as possible, even though not all the terms are used in all varieties of English, and some of them may be used in primarily literary contexts rather than everyday speech.

14. The word *odor* was not included in counts, but is shown here as a root form for quality terms such as *odorous* and *odoriferous*. In counting quality terms in both languages, we included all smell-related occurrences of the relevant roots, including conventional phrases such as *stink bug*, and proper names that include a smell term. For terms that have a homonym with a non-odor meaning (e.g., *rank* in English; or *pudyu*, a variant of *pu-tyu* 'be-NEG' in Ch'apalaa), these meanings were discarded either through hand-checking of examples, or through hand-counting odor versus non-odor meanings in a representative sample and applying this rate of occurrence to the total figure returned in automatic searches.

15. Cha'palaa smell quality terms appear more frequently than English smell quality terms, when comparing separately: English COCA corpus $t(14.07) = 2.05$, $p = .03$, Santa Barbara corpus $t(14.73) = 1.90$, $p = .035$, Rossi corpus $t(19.27) = 1.40$, $p = .08$, or all corpora combined $t(14.07) = 2.05$, $p = .03$. These comparisons are based on adjusted frequencies (i.e., counts divided by the total number of words in the corpus), since corpus sizes vary so greatly. Also, homogeneity of variance was not assumed for the statistics.

16. The higher rate of references to smell also has implications for socialization of Chachi children, who are thus more often exposed to complex olfactory language and olfaction-attuned behavior in their society.

17. These types of complex scenarios of relativity are more along the lines of what Whorf originally proposed than the simplistic versions targeted by claims of detractors that "the world looks the same in any language", in terms of McWhorter (2014). Previously Webster (2015:87) has pointed out the visual bias in this formulation by asking about ways that "the world doesn't sound the same in any language". We can ask whether the world smells the same in every language, and in light of languages like Cha'palaa, it would seem the answer is "no".

References

Ackerman, Diane. 1991. *A Natural History of the Senses*. New York: Vintage Books.

Adelaar, Willem F. H., and Pieter Muysken. 2004. *The Languages of the Andes*. Cambridge, UK: Cambridge University Press.

Aikhenvald, Alexandra Y. 2003. *Classifiers: A Typology of Noun Categorization Devices*. Oxford: Oxford University Press.

Aikhenvald, Alexandra Y., and Anne Storch. 2013. "Linguistic Expression of Perception and Cognition: A Typological Glimpse." In *Perception and Cognition in Language and Culture*, edited by Alexandra Y. Aikhenvald and Anne Storch, volume 3, 1–45. Leiden, the Netherlands: Brill. http://www.brill.com/perception-and-cognition-language-and-culture.

Amberber, Mengistu, Brett Baker, and Mark Harvey. 2010. *Complex Predicates: Cross-Linguistic Perspectives on Event Structure*. Cambridge, UK: Cambridge University Press.

Barkat-Defradas, Melissa, and Elisabeth Motte-Florac. 2016. *Words for Odours: Language Skills and Cultural Insights*. Newcastle upon Tyne, UK: Cambridge Scholars Publishing.

Barrett, Samuel Alfred. 1925. *The Cayapa Indians of Ecuador*. New York: Museum of the American Indian, Heye Foundation.

Baugh, Albert Croll, and Thomas Cable. 2012. *A History of the English Language*. New York: Pearson.

Beach, Lee Roy. 1964. "Cue Probabilism and Inference Behavior." *Psychological Monographs: General and Applied* 78(5–6): 1–20. https://doi.org/10.1037/h0093853.

Beek, Walter E. A. van van. 1992. "The Dirty Smith: Smell as a Social Frontier among the Kapsiki/Higi of North Cameroon and North-Eastern Nigeria." *Africa* 62(1): 38–58. https://doi.org/10.2307/1160063.

Berlin, Brent, and Paul Kay. 1969. *Basic Color Terms: Their Universality and Evolution*. Berkeley and Los Angeles: University of California Press.

Blench, Roger M., and Selbut R. Longtau. 1995. "Tarok Ophresiology: An Investigation into the Tarok Terminology of Odours." In *Issues in African Languages and Linguistics: Essays in Honour of Kay Williamson*, special issue of Nigerian Language Studies: 340–344. Aba (Nigeria): National Institute for Nigerian Languages.

Burenhult, Niclas, and Asifa Majid. 2011. "Olfaction in Aslian Ideology and Language." *The Senses and Society* 6(1): 19–29. https://doi.org/10.2752/174589311X12893982233597.

Bybee, Joan L. 2007. "From Usage to Grammar: The Mind's Response to Repetition." *Language* 82(4): 711–733. https://doi.org/10.1353/lan.2006.0186.

Cerrón-Palomino, Rodolfo. 1987. *Lingüística quechua* Cusco, Peru: Centro de Estudios Rurales Andinos "Bartolomé de las Casas."

Classen, Constance. 1997. "Foundations for an Anthropology of the Senses." *International Social Science Journal* 49(153): 401–412. https://doi.org/10.1111/j.1468-2451.1997.tb00032.x.

Classen, Constance, David Howes, and Anthony Synnott. 1994. *Aroma: The Cultural History of Smell*. London: Routledge.

Cole, Peter. 1985. *Imbabura Quechua*. Israeli Hebrew.

Conklin, Harold C. 1955. "Hanunóo Color Categories." *Southwestern Journal of Anthropology* 11(4): 339–344. https://doi.org/10.1086/soutjanth.11.4.3628909.

Corbett, Greville G. 1991. *Gender*. Cambridge, UK: Cambridge University Press.

Corbin, Alain. 1986. *The Foul and the Fragrant: Odor and the French Social Imagination*. Cambridge, MA: Harvard University Press.

Cox, Christopher. 2011. "Corpus Linguistics and Language Documentation: Challenges for Collaboration." In *Corpus-Based Studies in Language Use, Language Learning, and Language Documentation*, 239–264. Leiden, the Netherlands: Brill. https://doi.org/10.1163/9789401206884_013.

Crystal, David. 2003. *English as a Global Language*. Cambridge, UK: Cambridge University Press.

Curnow, Timothy Jowan, and Anthony J. Liddicoat. 1998. "The Barbacoan Languages of Colombia and Ecuador." *Anthropological Linguistics* 40(3): 384–408.

Dickinson, Connie. 2002. *Complex Predicates in Tsafiki*. Eugene: University of Oregon.

Dingemanse, Mark, and Simeon Floyd. 2014. "Conversation across Cultures." In *The Cambridge Handbook of Linguistic Anthropology*, edited by N. J. Enfield, Paul Kockelman, and Jack Sidnell, 447–480. Cambridge, UK: Cambridge University Press.

Du Bois, John W., Wallace L. Chafe, Charles Meyer, Sandra A. Thompson, Robert Englebretson, Nii Martey. 2000–2005. *Santa Barbara Corpus of Spoken American English*, Parts 1–4. Philadelphia: Linguistic Data Consortium.

Floyd, Simeon. 2009. "Nexos Históricos, Gramaticales Y Culturales de Los Números En Cha'palaa." In *Proceedings from CILLA IV*. Austin, TX.

———. 2010. "Discourse Forms and Social Categorization in Cha'palaa." PhD dissertation, University of Texas at Austin.

———. 2014a. "'We'as Social Categorization in Cha'palaa, a Language of Ecuador." *Constructing Collectivity:'We'across Languages and Contexts* 239. http://books.google.nl/books?hl=nl&lr=&id=63raAgAAQBAJ&oi=fnd&pg=PT145&dq=simeon+floyd+social+categorization&ots=LJDBbXFWDv&sig=9D0nwebQpcuzCZmPX0Y_jco4iGk.

———. 2014b. "Four Types of Reduplication in the Cha'palaa Language of Ecuador." In *Reduplication in Indigenous Languages of South America*, edited by Hein Voort and Gale Goodwin Gómez, 77–113. Leiden, the Netherlands: Brill.

———. Forthcoming. "Egophoricity and Argument Structure in Cha'palaa." In *Egophoricity*, edited by Simeon Floyd, Lila San Roque, and Elisabeth Norcliffe. Typological Studies in Language. John Benjamins.

Gippert, Jost, Nikolaus P. Himmelmann, and Ulrike Mosel. 2006. *Essentials of Language Documentation*. Berlin: De Gruyter Mouton.

Goodwin, Charles. 1997. "The Blackness of Black: Color Categories as Situated Practice." In *Discourse, Tools and Reasoning*, edited by Lauren B. Resnick, Roger Säljö, Clotilde Pontecorvo, and Barbara Burge, 111–40. NATO ASI Series 160. Berlin Heidelberg: Springer. https://doi.org/10.1007/978-3-662-03362-3_6.

Henrich, Joseph, Steven J. Heine, and Ara Norenzayan. 2010. "The Weirdest People in the World?" *Behavioral and Brain Sciences* 33(2–3): 61–83. https://doi.org/10.1017/S0140525X0999152X.

Hombert, Jean Marie. 1992. "Terminologie des odeurs dans quelques langues du Gabon." *Pholia* 7: 61–65.

Hopper, Paul J. 1996. "Some Recent Trends in Grammaticalization." *Annual Review of Anthropology* 25: 217–236.

Hopper, Paul J., and Elizabeth Closs Traugott. 2003. *Grammaticalization*, 2nd ed. Cambridge, UK: Cambridge University Press.

Howes, David. 1991. "Sense and Non-Sense in Contemporary Ethno/Graphic Theory and Practice." *Culture* 11: 51–62.

———. 2003. *Sensual Relations: Engaging the Senses in Culture and Social Theory*. Ann Arbor: University of Michigan Press.

———. 2005. "Introduction: Empire of the Senses." In *Empire of the Senses: The Sensual Culture Reader*, edited by David Howes, 1–22. London: Bloomsbury Academic.

Howes, David, and Sarah Pink. 2010. "The Future of Sensory Anthropology/The Anthropology of the Senses." *Social Anthropology* 18(3): 331–340. https://doi.org/10.1111/j.1469-8676.2010.00119_1.x.

Ingold, Tim. 2000. *The Perception of the Environment: Essays on Livelihood*. Hove, UK: Psychology Press.

———. 2011. "Worlds of Sense and Sensing the World: A Response to Sarah Pink and David Howes." *Social Anthropology* 19(3): 313–317. https://doi.org/10.1111/j.1469-8676.2011.00163.x.

Kay, Paul, Brent Berlin, and Luisa Maffi. 2011. *The World Color Survey*. Stanford, CA: Center for the Study of Language and Information Publications.

Kay, Paul, and Terry Regier. 2003. "Resolving the Question of Color Naming Universals." *Proceedings of the National Academy of Sciences* 100(15): 9085–9089. https://doi.org/10.1073/pnas.1532837100.

Kung, Susan Smythe. 2007. "A Descriptive Grammar of Huehuetla Tepehua." Thesis. https://repositories.lib.utexas.edu/handle/2152/3137.

Lawless, H., and T. Engen. 1977. "Associations to Odors: Interference, Mnemonics, and Verbal Labeling." *Learning and Memory* 3(1): 52–59.

Levinson, Stephen C. 2000. "Yélî Dnye and the Theory of Basic Color Terms." *Journal of Linguistic Anthropology* 10(1): 3–55. https://doi.org/10.1525/jlin.2000.10.1.3.

Llitjós, Ariadna Font, Roberto Aranovich, and Lori Levin. 2005. "Building Machine Translation Systems for Indigenous Languages." In *Second Conference on the Indigenous Languages of Latin America (CILLA II), Texas, USA*. http://www.cs.cmu.edu/~aria/Papers/FontAranovich_CILLA2_mapuche_quechua(2).pdf.

Lucy, John A. 1997. "The Linguistics of 'Color.'" In *Color Categories in Thought and Language*, edited by C. L. Hardin and Luisa Maffi, 320–346. Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/cbo9780511519819.015.

Majid, Asifa. 2015a. "Comparing Lexicons Cross-Linguistically." In *The Oxford Handbook of the Word*, edited by John R. Taylor. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199641604.013.020.

———. 2015b. "Cultural Factors Shape Olfactory Language." *Trends in Cognitive Sciences* 19 (11): 629–630. https://doi.org/10.1016/j.tics.2015.06.009.

Majid, Asifa, and Niclas Burenhult. 2014. "Odors Are Expressible in Language, as Long as You Speak the Right Language." *Cognition* 130(2): 266–270.

Majid, Asifa, and Nicole Kruspe. 2018. "Hunter-Gatherer Olfaction Is Special." *Current Biology* 28: 409–413. https://doi.org/10.1016/j.cub.2017.12.014.

Majid, Asifa, and Stephen C. Levinson. 2011. "The Senses in Language and Culture." *The Senses and Society* 6(1): 5–18. https://doi.org/10.2752/174589311X12893982233551.

Martin, Laura. 1986. "'Eskimo Words for Snow': A Case Study in the Genesis and Decay of an Anthropological Example." *American Anthropologist* 88(2): 418–423.

McEnery, Tony, Paul Baker, and Lou Burnard. 2000. "Corpus Resources and Minority Language Engineering." In *LREC*. Citeseer. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.679.3333&rep=rep1&type=pdf.

McEnery, Tony, and Andrew Hardie. 2011. *Corpus Linguistics: Method, Theory and Practice*. Cambridge, UK: Cambridge University Press.

McGann, John P. 2017. "Poor Human Olfaction Is a 19th-Century Myth." *Science* 356(6338): eaam7263. doi:10.1126/science.aam7263.

McWhorter, John H. 2014. *The Language Hoax: Why the World Looks the Same in Any Language*. Oxford: Oxford University Press.

Olofsson, Jonas K., and Jay A. Gottfried. 2015. "The Muted Sense: Neurocognitive Limitations of Olfactory Language." *Trends in Cognitive Sciences* 19(6): 314–321. https://doi.org/10.1016/j.tics.2015.04.007.

O'Meara, Carolyn, and Asifa Majid. 2016. "How Changing Lifestyles Impact Seri Smellscapes and Smell Language." *Anthropological Linguistics*. 58(2): 107–131.

Pullum, Geoffrey K. 1991. *The Great Eskimo Vocabulary Hoax and Other Irreverent Essays on the Study of Language*. Chicago: University of Chicago Press.

Quirk, Randolph. 1980. *A Grammar of Contemporary English*. London: Longman.

Reed, Stephen K. 1972. "Pattern Recognition and Categorization." *Cognitive Psychology* 3(3): 382–407. https://doi.org/10.1016/0010-0285(72)90014-X.

Regier, Terry, Alexandra Carstensen, and Charles Kemp. 2016. "Languages Support Efficient Communication about the Environment: Words for Snow Revisited." *PLoS ONE* 11(4): e0151138. https://doi.org/10.1371/journal.pone.0151138.

Rosch, Eleanor, Carolyn B. Mervis, Wayne D. Gray, David M. Johnson, and Penny Boyes-Braem. 1976. "Basic Objects in Natural Categories." *Cognitive Psychology* 8(3): 382–439. https://doi.org/10.1016/0010-0285(76)90013-X.

Rossi, Giovanni, Kobin H. Kendrick. 2013. *The Rossi Corpus of Conversational English*. Nijmegen, the Netherlands: Max Planck Institute for Psycholinguistics.

Salomon, Frank. 2007. *Native Lords of Quito in the Age of the Incas: The Political Economy of North Andean Chiefdoms*. Cambridge, UK: Cambridge University Press.

Samvelian, Pollet, and Léa Nash, eds. 2015. *Approaches to Complex Predicates*. Leiden, the Netherlands: Brill. http://booksandjournals.brillonline.com/content/books/9789004307094.

San Roque, Lila, Kobin H. Kendrick, Elisabeth Norcliffe, Penelope Brown, Rebecca Defina, Mark Dingemanse, Tyko Dirksmeyer, et al. 2015. "Vision Verbs Dominate in Conversation across Cultures, but the Ranking of Non-Visual Verbs Varies." *Cognitive Linguistics* 26(1). https://doi.org/10.1515/cog-2014-0089.

Scannell, Kevin P. 2007. "The Crúbadán Project: Corpus Building for Under-Resourced Languages." In *Building and Exploring Web Corpora: Proceedings of the 3rd Web as Corpus Workshop*, 4: 5–15. http://books.google.com/books?hl=en&lr=&id=EgeERqdkn20C&oi=fnd &pg=PA5&dq=info:vgfzKFv7mZ4J:scholar.google.com&ots=v9Nou9-u-r&sig=RzoEtf hj2uztncqcL-93KkjZzIA.

Seifart, Frank. 2010. "Nominal Classification: Nominal Classification." *Language and Linguistics Compass* 4(8): 719–736. https://doi.org/10.1111/j.1749-818X.2010.00194.x.

Seifart, Frank, and Doris L. Payne. 2007. "Nominal Classification in the North West Amazon: Issues in Areal Diffusion and Typological Characterization." *International Journal of American Linguistics* 73(4): 381–387. https://doi.org/10.1086/523770.

Sene-Mongaba, Bienvenu. 2015. "The Making of Lingala Corpus: An Under-Resourced Language and the Internet." *Procedia - Social and Behavioral Sciences*, Current Work in Corpus Linguistics: Working with Traditionally-conceived Corpora and Beyond. Selected Papers from the 7th International Conference on Corpus Linguistics (CILC2015), 198 (July): 442–450. https://doi.org/10.1016/j.sbspro.2015.07.464.

Senft, Gunter. 2000. *Systems of Nominal Classification*. Cambridge, UK: Cambridge University Press.

Shepard, Glenn Harvey Jr. 1999. *Pharmacognosy and the Senses in Two Amazonian Societies*. University of California, Berkeley, with the University of California, San Francisco.

Slobin, Dan I. 1987. "Thinking for Speaking." *Annual Meeting of the Berkeley Linguistics Society* 13: 435–445. https://doi.org/10.3765/bls.v13i0.1826.

Sloetjes, Han, and Olaf Seibert. 2016. "Measuring by Marking: the Multimedia Annotation Tool ELAN." In *Measuring Behavior 2016, 10th International Conference on Methods and Techniques in Behavioral Research*, edited by A. Spink, G. Riedel, L. Zhou, L. Teekens, R. Albatal, and C. Gurrin, 492–95.

Sorokowska, Agnieszka, Piotr Sorokowski, Thomas Hummel, and Tomas Huanca. 2013. "Olfaction and Environment: Tsimane' of Bolivian Rainforest Have Lower Threshold of Odor Detection Than Industrialized German People." *PLOS ONE* 8(7): e69203. https://doi. org/10.1371/journal.pone.0069203.

Sperber, Dan. 1975. *Rethinking Symbolism*. Cambridge, UK: Cambridge University Press Archive.

Stivers, T., N. J Enfield, P. Brown, C. Englert, M. Hayashi, T. Heinemann, G. Hoymann, et al. 2009. "Universals and Cultural Variation in Turn-Taking in Conversation." *Proceedings of the National Academy of Sciences* 106(26): 10587.

Stoddart, David Michael. 1990. *The Scented Ape: The Biology and Culture of Human Odour*. Cambridge, UK: Cambridge University Press.

Stoller, Paul. 1989. *Fusion of the Worlds: An Ethnography of Possession Among the Songhay of Niger*. Chicago: University of Chicago Press.

Storch, Anne, and Rainer Vossen. 2007. "Odours and Colours in Nilotic: Comparative Case Studies." In *Advances in Nilo-Saharan Linguistics: Proceedings of the 8th Nilo-Saharan Linguistics Colloquium, University of Hamburg, August 22–25, 2001*, edited by Doris L. Payne and Mechthild Reh, 223–240. Nilo-Saharan, v. 22. Cologne: Köppe.

Surrallés, Alexandre. 2016. "On Contrastive Perception and Ineffability: Assessing Sensory Experience without Colour Terms in an Amazonian Society." *Journal of the Royal Anthropological Institute* 22(4): 962–979. https://doi.org/10.1111/1467-9655.12499.

Torero, Alfredo. 1964. *Los dialectos quechuas*. Lima: Map.

Tufvesson, Sylvia. 2011. "Analogy-Making in the Semai Sensory World." *The Senses and Society* 6(1): 86–95. https://doi.org/10.2752/174589311X12893982233876.

Viberg, Åke. 1983. "The Verbs of Perception: A Typological Study." *Linguistics* 21(1). https:// doi.org/10.1515/ling.1983.21.1.123.

Viveiros de Castro, Eduardo. 2011. "Perspectival Anthropology and the Method of Controlled Equivocation." *Tipití: Journal of the Society for the Anthropology of Lowland South America* 2(1). http://digitalcommons.trinity.edu/tipiti/vol2/iss1/1.

Webster, Anthony K. 2015. "Why the World Doesn't Sound the Same in Any Language and Why That Might Matter: A Review of The Language Hoax: Why the World Looks the Same

in Any Language, by John H. McWhorter. Oxford: Oxford University Press. 2014. [Xx + 182 Pages]." *Journal of Linguistic Anthropology* 25(1): 87–93. https://doi.org/10.1111/jola.12068.

Wierzbicka, Anna. 2005. "There Are No 'Color Universals' but There Are Universals of Visual Semantics." *Anthropological Linguistics* 47(2): 217–244. https://doi.org/10.2307/25132327.

Wilson, Donald A., and Richard J. Stevenson. 2006. *Learning to Smell: Olfactory Perception from Neurobiology to Behavior*. Baltimore: Johns Hopkins University Press.

Winter, Bodo. 2016. "The Sensory Structure of the English Lexicon." *eScholarship*, January. http://escholarship.org/uc/item/885849k9.

Wnuk, Ewelina, and Asifa Majid. 2014. "Revisiting the Limits of Language: The Odor Lexicon of Maniq." *Cognition* 131(1): 125–138. https://doi.org/10.1016/j.cognition.2013.12.008.

Yeshurun, Yaara, and Noam Sobel. 2010. "An Odor Is Not Worth a Thousand Words: From Multidimensional Odors to Unidimensional Odor Objects." *Annual Review of Psychology* 61: 219–241. https://doi.org/10.1146/annurev.psych.60.110707.163639.