Linguistica Silesiana nr 441, 2023
ISSN 0208-4228
DOI: 10.24425/linsi.2023.144824

ALEXANDER ANDRASON
Living Tongues Institute for Endangered Languages (Salem, USA)
University of Cape Town (South Africa)
aleksand@hi.is

THE NON-ARBITRARINESS OF SOME CONATIVE CALLS USED TO CHASE ANIMALS *

The present article demonstrates that languages tend to contain dispersals – a subtype of conative calls used to chase animals – that are built around voiceless sibilants. This tendency is both quantitative (i.e., voiceless-sibilant dispersals are common across languages and in a single language) and qualitative (i.e., sibilants contribute very significantly to the phonetic substance of such dispersals). This fact, together with a range of formal similarities exhibited by voiceless-sibilant dispersals encapsulated by the pattern [kI/Uʃ] suggests that the presence of voiceless sibilants in dispersals is not arbitrary. Overall, voiceless-sibilant dispersals tend to comply with the general phonetic profile associated with the prototype of CACs and dispersals, postulated recently in scholarship, thus corroborating the validity of this prototype.

Keywords: interjections, conative animal calls, dispersals, typology

1. Introduction

Conative animal calls (CACs) are lexicalized constructions – or synthetic form-meaning combinations (Goldberg 2003; Fried and Östman 2004; Hoffmann and Trousdale 2013) – that exhibit a fully entrenched "directive-to-animal function" (Andrason and Karani 2021: 34-35: Andrason 2022: 27-28). That is, CACs convey "requests, wishes, desires, demands, or orders" and, contrary to many other linguistic categories, they are primarily addressed to animals (Andrason and Karani 2021: 33). To put it differently, CACs constitute a particular subgroup of conatives (Ameka 1992a),[1] namely the one with which

* I would like to thank Professors Bernd Heine and Mirosław Bańko for their very valuable comments on the original draft of this paper.
[1] Conatives are also referred to as 'volitives' (Wierzbicka 2003) or 'directives' (Heine 2023).

humans influence the behavior of other species rather than the behavior of the members of their own species (Andrason and Karani 2021: 3-4).

Thus far, to my knowledge, CACs have only been studied in a thorough and systematic manner in six language systems: in a Tamazight (Berber) dialect spoken in Ayt Hadiddu in Morocco (Bynon 1976), in Polish and some other West-Slavonic languages (Indo-European) spoken in Central Europe (Siatkowska 1976), in Zargulla (Omotic) spoken in Ethiopia (Amha 2013), in Chuvash (Turkic) spoken in the Chuvash Republic in Russia (Denisova and Sergeev 2015), in Arusa Maasai (Nilotic) spoken in Tanzania (Andrason and Karani 2021), and in Xhosa (Bantu) spoken in South Africa (Andrason 2022).[2] As a result, our knowledge of crosslinguistic tendencies governing the form and meaning of CACs is severely limited – CACs no doubt require typological research (Poyatos 1993: 442; 2002: 187; Andrason and Karani 2021: 4-5).[3]

Although some shortcomings in the typology of CACs have recently been rectified – see Andrason and Karani (2021), Andrason (2022), and Heine (2023), who have proposed the crosslinguistic prototype of a CAC and distinguished or tested a number of prototypical features, whether semantic, pragmatic, phonetic, morphological, or syntactic – several aspects of the behavior of CACs in the languages of the world remain unknown or controversial. One of them is the issue of the arbitrariness or non-arbitrariness of CACs (Poyatos 1993: 443-444; 2002: 187). As noted by Bynon (1976: 62) almost fifty years ago, "certain apparently arbitrary calls, in particular scares to both wild and domestic animals but also certain other domestic animal calls, may contain an acoustic component having an inherently attracting or repelling effect upon the species to which it is addressed". The present research aims to contribute to the scholarly understanding of the relationship between the form of CACs and their function by studying commonalities in the phonetic substance exhibited by one of the two CAC types mentioned by Bynon (ibid.) as potentially non-arbitrary – dispersals, i.e., CACs with which humans repel, chase, or drive away animals (Amha 2013; Andrason and Karani 2021).[4]

During my studies dedicated to CACs in Arusa Maasai (Andrason and Karani 2021), Xhosa (Andrason 2022), and Polish, I have indeed noticed certain phonetic commonalities in dispersals, specifically, a pervasive use of voiceless

[2] As two of these studies, namely Siatkowska (1976) and Denisova and Sergeev (2015) are not written in English but Polish and Russian, respectively, they have escaped the attention of various linguists dealing with CACs (see Poyatos 1993; 2002; Aikenvald 2010) including myself (Andrason and Karani 2021).

[3] For a more thorough review of research on CACs, both language-specific and typological, consult Andrason and Karani (2021: 3-5).

[4] Other types of CACs are 'summonses' which are used to "make the animal move closer to the speaker" (Andrason and Karani 2021: 9) and 'kinetics' (also referred to as 'directives') used to "initiate, sustain, terminate, or modify the motion of an animal" (ibid. 10). In this article, I reserve the term 'directive' to the function exhibited by all CACs.

sibilants.[5] Given the absence of any genetic and areal relationship between these three languages and in light of a similar tendency emerging from the available evidence from Ayt Hadiddu (Bynon 1976), Chuvash (Denisova and Sergeev 2015), Ewe (Ameka 1992b), Lithuanian (Ambrazas et al. 2006), Matses (Fleck 2003), Zargulla (Amha 2013), and West-Slavonic languages (Siatkowska 1976), which I had also observed in the course of my previous research activities, this commonality did not seem accidental to me. I reckoned that it was rather due to the non-arbitrary foundation of these types of CACs. As a result, I decided to examine a more specific hypothesis and verify the extent to which the entanglement of dispersals and voiceless sibilants is crosslinguistically pervasive, i.e.: Do languages tend to contain dispersals that are built around voiceless sibilants? The more visible – both quantitatively (i.e., by being frequent across languages and in a single language) and qualitatively (i.e., by contributing more significantly to the phonetic substance of a token)[6] – voiceless sibilants are in dispersals, the more non-arbitrary (these types of) dispersals arguably are. Additionally, I chose to test all voiceless-sibilant dispersals that I would collect for their compliance with the general phonetic profile associated with the prototype of CACs. Given the correlation of cognitive salience with the CAC prototype suggested recently in scholarship (Andrason and Karani 2021: 8), if voiceless-sibilant dispersals are non-arbitrary, they should exhibit (most of) the prototypical features. Should this be true, the validity of the prototype would, in turn, be confirmed. Should this not be true, certain aspects of a prototypical CAC would need to be revised.[7]

As can be inferred from the previous paragraphs, my research will be developed within a typologically informed prototype-driven approach to the category of CACs (Andrason and Karani 2021; Andrason 2022). Additionally, the description of the phonetic substance of CACs will be carried out within a non-formal approach to the sound system, which is typical of Basic Linguistic Theory (Dryer 2006; Dixon 2010) and which also permeates most typological studies.

In order to achieve my research objectives, this article is structured in the following manner: in Section 2, I present the framework underlying my study; in Section 3, I introduce typological evidence; in Section 4, I evaluate this evidence and suggest potential implications for the general theory of dispersals and CACs; in Section 5, I conclude my paper.

---

[5] For the definition of a sibilant see Section 2.

[6] I use the term 'qualitative' rather loosely. Of course, my qualitative analysis also draws on quantitative evidence, i.e., the number of sibilants and the other phones in a word. However, the goal of this qualitative analysis is to establish not how common voiceless-sibilant dispersals are, but how important/relevant/significant voiceless sibilants are in a dispersal.

[7] My method is thus similar to the method adopted in the studies dedicated to the non-arbitrariness of the encodings of the concept of 'frog' conducted by Berlin (1992) and Hays (1994) In these studies, the prevalence of [r]/[l] and [g] has been examined.

2. Framework

As I mentioned in the previous section, my study is anchored in two theoretical frames of reference: a typologically informed, prototype-driven approach to the category of CACs and a non-formal approach to the sound system of languages typical of Basic Linguistic Theory and widely embraced by typologists.

The methodological backbone of my research is provided by a prototype-driven approach to linguistic categorization, which, from a more general perspective, characterizes cognitive linguistics (Taylor 1995; Evans and Green 2006; Janda 2015) – a scientific paradigm to which I subscribe. To be exact, following Andrason and Karani (2021), as well as Andrason (2022) and Heine (2023), I view the category of CACs as a radial network with prototype effects. The CAC category is organized around an ideal representative – the prototype. The prototype represents the conceptual nucleus of the category and is used as a tool to measure the categorial status of all the other members. Members that match the prototype closely are canonical and occupy the center of the category. In contrast, members that match the prototype partially or minimally are non-canonical and populate the category's periphery. Importantly, although the membership status of members is not uniform, ranging from more to less canonical, all such members belong to the category. Membership is thus not a binary issue of inclusion or exclusion but rather a matter of degree (Andrason and Karani 2021: 5-8; see also Janda 2015). Therefore, the semantic-pragmatic definition of CACs provided at the beginning of this article does not imply that analytical CACs or CACs in which a directive-to-animal function is entrenched to a lesser extent, do not belong to the CAC category. Their categorial membership is intact – they are merely less canonical and more peripheral.

Given the categorization approach chosen, the critical notion for the category of CACs is the prototype. This prototype – recently proposed by Andrason and Karani (2021) and subsequently reexamined by Andrason (2022) and Heine (2023) – is defined cumulatively as a set of properties. These prototypical properties, which are varied and concern semantics, pragmatics, syntax, morphology, and phonetics, have been posited given their crosslinguistic pervasiveness and/or cognitive salience, as well as diachronic convergence and presence in spontaneous coinage episodes (Andrason and Karani 2021: 5-8). For the present study, only the phonetic properties associated with the prototype are important since I analyze the relationship between the meaning of CACs (specifically those classified as dispersals) and their phonetic substance.

The following phonetic properties are viewed as prototypical in literature:

(a) A prototypical CAC "exhibits a consonantal nature" – it heavily draws on consonantal material or consists entirely of consonants (Andrason and Karani 2021: 34).

(b) A prototypical CAC contains extra-systematic sounds. This extra-systematicity can be of two types: sounds used are "foreign to human language in general" and do not feature in the IPA alphabet (Andrason and Karani 2021: 34) or they are merely "foreign to the language in which particular CACs are found" (ibid.). The most characteristic representatives of the former type are whistles and kissing sounds; the most characteristic for the latter type are clicks (ibid.). Although consonantal and/or vocalic extra-systematicity is the feature of the prototype, it need not be frequent at the category level (Andrason 2022: 43). That is, what distinguishes real-world CACs from words or constructions of other classes, is "the *grammaticality* of such extra-systematic phones, i.e., the fact that they *can* be attested", rather than the *commonness* of extra-systematic phones in CACs across languages (ibid.; emphasis is mine).

(c) A prototypical CAC is extra-systematic from a phonotactic, suprasegmental, and phonation-related perspective (Andrason 2022: 44). Contrary to the extra-systematicity of phones mentioned above, these extra-systematic properties are clearly visible at a category level, thus being common in real-world CACs. That is, CACs are highly susceptible to the so-called extensions, i.e., "prolongation, replication, and repetition" and modulations, i.e., "intensity, loudness, [and marked] rate of delivery and intonation" (Andrason and Karani 2021: 34).

(d) The abovementioned phonotactic, suprasegmental, and phonation-related properties are correlated with the meaning of CACs. Accordingly, a prototypical dispersal is short, i.e., monosyllabic (Andrason 2022), realized with "articulatory speed (i.e., short rate of production)" (Andrason and Karani 2021: 35), and pronounced with "a raised voice" (ibid.), i.e., loudly and aggressively (ibid. 21, 35). Inversely, multisyllabicity, replications, and "friendly intonation" (ibid.) are rare, while repetitions, although relatively common, are optional.

While the typologically informed and prototype-driven approach to CACs will guide my research by suggesting phonetic features that should be examined, the description of the phonetic material collected itself will be achieved by making use of the terminological apparatus provided by Basic Linguistic Theory (Dryer 2006: 210-212; Dixon 2010: 264-288). Basic Linguistic Theory is a common – probably the most common – descriptive framework used in linguistic typology (Dryer 2006: 211; Dixon 2010). It follows a fundamental tenet that characterizes all non-formal approaches to language, which Goldberg (2003: 219) eloquently labelled as a "what you see is what you get" principle. This means that, as far as sound system is concerned, no 'deep levels' of analysis or "empty elements" are necessary (ibid.)

To avoid an excessive dependency on the theory underlying the analysis of sounds – the phenomenon virtually unavoidable in phonology where models for the same language are often radically different – my description will concern phonetics.[8] I will thus analyze the sound properties of the voiceless-sibilant dispersals in terms of their articulation and phonation rather than in terms of the contrast they yield in the entire language system (Dixon 2010: 338). My main empirical task will consist of transcribing the consonantal and vocalic material of the collected CACs with IPA symbols (see Section 3). These transcriptions will subsequently allow me to determine the types of sibilants used, the phonotactics and syllable structure of the dispersals collected, especially the features related to prolongation, replication, and repetition, and the properties of the other phones present in them. Additionally, I will refer to the phonation-related features, such as intensity, loudness, and rate of delivery (see Section 4).

As is evident from the discussion above, the crucial phonetic elements in my study are sibilants. Sibilants constitute a subtype of fricatives that are produced in two regions in the oral cavity: dental and alveolar (Ladefoged and Maddieson 1996: 145). The main source of sibilants is a "turbulent airstream" (ibid.). This airstream emerges "when the jet of air created by the dental or alveolar constriction strike the teeth, which form an obstacle downstream from the constriction itself" (ibid.). The typical voiceless sibilants, the most relevant types of sibilants for my research, are [s] (alveolar), [ʃ] (partially palatalized postalveolar or palato-alveolar), [ɕ] (alveolo-palatal), and [ʂ] (retroflex; also including a laminal flat postalveolar variant [s̻]). Additionally, I will include in my review lexemes containing affricates that have a sibilant release, i.e., [t͡s], [t͡ʃ], [t͡ɕ], and [t͡ʂ]. All these "plain" sibilant phones may also exhibit secondary articulation. To be exact, they can be labio-velarized [ʷ], labialized (and whistled) [ʱ], palatalized [ʲ], velarized [ˠ], glottalized / ejective [ʼ], and pharyngealized [ˤ].

3. Evidence

The evidence presented in this section has been collected in a heterogenous manner. This includes physical and remote fieldwork activities carried out in 2021, interviews with native speakers, collaboration with linguists around the globe, and reviews of published works (articles, grammars, and dictionaries).

[8] Certainly, phonetic analyses of sounds (including sibilants) can also differ. Nevertheless, such alternative descriptions are generally less conflicting than divergent phonemic models, and the selection of one of the phonetic descriptions has much less critical bearings on the result of a study. In my empirical study (see Table 1 in Section 3), the relevant alternative phonetic descriptions will be acknowledged.

Overall, while searching for dispersals that would contain voiceless sibilants, I have studied (the categories of) CACs in 79 languages. Although I did not follow any randomized sampling procedure but instead took into consideration selective criteria when choosing the languages for my study, the languages included in my research are highly varied, both genetically and geographically.[9] This diversity, in turn, strengthens the crosslinguistic validity of potential observed generalizations. To be exact, the languages included in my study belong to 15 distinct linguistic families or realms: Indo-European (32 varieties from the Baltic, Germanic, Hellenic Indo-Iranian, Romance, and Slavonic branches), Niger-Congo (18 varieties from the Adamawa, Bantu, Cangin, Kwa, and Volta-Congo branches), Afroasiatic (9 varieties from the Berber, Chadic, Cushitic, Omotic, and Semitic branches), Turkic (4 varieties from the Oghur, Kipchak, and Oghuz branches), Uralic (3 varieties from the Finno-Ugric branch), Austronesian (2 varieties from the Malayo-Polynesian branch), and Nilotic (2 varieties), as well as one variety from the Austroasiatic (the Vietic branch), Dravidian, Japonic, Koreanic, Mongolic, Panoan, Sino-Tibetan, and Uto-Aztecan families each. Additionally, one variety represents language isolates. Similar diversity concerns the areas in which the languages reviewed are spoken – their speakers inhabit Africa, the Americas, Asia, and Europe.

The table below summarizes the collected evidence. It presents the respective languages ordered alphabetically, their phylogenetic classification, the voiceless-sibilant dispersals that I have found, their IPA transcription, the particular sibilant used, and the source of my data.

Table 1: Voiceless-sibilant dispersals in the languages of the world

Language	Family	Lexeme(s)	IPA	Sibilant	Source
Afrikaans	Germanic (West) (Indo-European)	*shoe*	[ʃu]	[ʃ]	own data
Akan	Kwa (Niger-Congo)	*su*(*u*)	[su(:)]	[s]	own data
Arabic	Semitic (Afroasiatic)	غيس	[ʁis]	[s]	Abdulla and Talib (2009) own data
		هس	[hus:]	[s]	
		هيش	[hiʃ]	[s]	

[9] The inclusion of a language into my study was motivated by one of the following criteria: (a) I speak that variety or understand it; (b) I was able to involve the language experts and/or speakers of that variety in my research; or (c) the CACs of that variety have been described or mentioned in previous publications.

Table 1. cont.

Language	Family	Lexeme(s)	IPA	Sibilant	Source
Arusa	Maasai (Nilotic)	*sh*	[ʃ]	[ʃ]	Andrason and Karani (2021)
		shsh	[ʃː]	[ʃ]	
		s(s)	[s(ː)]	[s]	
		ssee	[sːeː]	[s]	
		ssuk	[sːuk]	[s]	
		ssek	[sːek]	[s]	
Ayt Hadiddu	Berber (Tamazight) (Afroasiatic)	*ṣab*	[sˤab]	[sˤ]	Bynon (1976)
		šta-ddu	[ʃtadːu]	[ʃ]	
		mšši	[mʃːi]	[ʃ]	
		xišš	[xiʃː]	[ʃ]	
		sar(r)	[sar(ː)]	[s]	
		hušš	[huʃː]	[ʃ]	
		kkušš	[kuʃː][10]	[ʃ]	
		kkšš	[kʃː]	[ʃ]	
Basque	Isolate	*xapi*	[ʃapi]	[ʃ]	etxepare.eus euskadi.eus
		sapi	[sapi]	[s]	
		gatxit	[gat͡ʃit]	[t͡ʃ]	
		gatx	[gat͡ʃ]	[t͡ʃ]	
		gats	[gat͡s]	[t͡s]	
		gaksss	[gaksː]	[s]	
		xipist	[ʃapist]	[ʃ] and [s]	
Bulgarian	Slavonic (South) (Indo-European)	*къш*	[kɤʃ]	[ʃ]	own data
Buryat	Central Mongolic (Mongolic)	*hээш*	[hɛːʃ]	[ʃ]	own data
Catalan	Romance (Occitan) (Indo-European)	*sap*	[sap]	[s]	Alcover and de Borja Moll (1969) diccionari.cat
		çapi	[sapi]	[s]	
		xo	[ʃo]	[ʃ]	
Changana	Bantu (S) (Niger-Congo)	*svi*	[sᵸi]	[sᵸ]	own data
		suka	[suka]	[s]	

[10] Geminated stops in a word-initial position exhibit a "delayed plosion with increase of tension" (Bynon 1976: 43).

Table 1. cont.

Language	Family	Lexeme(s)	IPA	Sibilant	Source
Chuvash	Oghur (Turkic)	*кăш*	[kʃ̌]	[ʃ]	Denisova and Sergeev (2015)
		кш(ш)а	[kʃ(:)a]	[ʃ]	
		кăшша	[kʃ̌(:)a]	[ʃ]	
		кшах	[kʃaχ]	[ʃ]	
		кшуй	[kʃuj]	[ʃ]	
		кш	[kʃ]	[ʃ]	
		хас	[χas]	[s]	
		прис	[pris]	[s]	
		качча	[kat͡ɕ(:)a]	[t͡ɕ]	
		кĕç	[kɔɛ]	[ɛ]	
		кçе	[kɛɛ]	[ɛ]	
		кăчăш	[kʃ̌t͡ɕʃ̌]	[t͡ɕ] and [ʃ]	
		путреш	[putrɛʃ]	[ʃ]	
		путряш	[putrjaʃ]	[ʃ]	
Cinyungwe	Bantu (N) (Niger-Congo)	*svi*	[sᵠi]	[sᵠ]	own data
		cika	[ʃika]	[ʃ]	
		sapi	[sapi]	[s]	
		coka	[ʃoka]	[ʃ]	
Czech	Slavonic (West) (Indo-European)	*kšá*	[kʃa]	[ʃ]	slovnik.seznam.cz ssjc.ujc.cas.cz own data
		kš(š)	[kʃ(:)]	[ʃ]	
		kšc	[kʃt͡s]	[ʃ] and [t͡s]	
Danish	Germanic (North) (Indo-European)	*ps(s)t*	[bs(:)d]	[s]	ordnet.dk/ddo own data
		pist	[pisd]	[s]	
Dholuo	West Nilotic (Nilotic)	*siauuu*	[sjau:]	[s]	own data
		ssch	[t͡ʃ:]	[t͡ʃ]	
		ss	[s:]	[s]	
Dutch	Germanic (West) (Indo-European)	*shoo*	[ʃu:]	[ʃ]	woorden.org own data
		ksj	[ksʲ]/[kʃ]	[sʲ]/[ʃ]	
		ksjt	[ksʲt]/[kʃt]	[sʲ]/[ʃ]	
		ks(s)t	[ks(:)t]	[s]	
Dzə (Jenjo)	Adamawa (Niger-Congo)	*sh*	[ʃ]	[ʃ]	own data
		su	[su]	[s]	

Table 1. cont.

Language	Family	Lexeme(s)	IPA	Sibilant	Source
		ss	[sː]	[s]	
		ʔⁱsɨ̀	[ʔⁱsɨ]	[s]	
English	Germanic (West) (Indo-European)	*shoo*	[ʃuː]	[ʃ]	Poyatos (1993; 2002) own data
		hsst	[sːt]	[s]	
		sk-sk$^{(n)}$	[sk.sk]	[s]	
Estonian	Finno-Ugric (Balto-Finnic) (Uralic)	*kõss*	[kɤsː]	[s]	eki.ee/dict/ekss/ eki.ee/EN/dictionaries
		ust	[ust]	[s]	
		uts	[uts]	[t͡s]	
Ewe	Volta-Congo (Niger-Congo)	*sɔ̃́*	[sɔ̃]	[s]	Ameka (1992b)
		súi	[sui]	[s]	
Finnish	Finno-Ugric (Balto-Finnic) (Uralic)	*hus*	[hus]	[s]	kotus.fi
French	Romance (Gallo) (Indo-European)	*(h)oust(e)*	[ust]	[s]	larousse.fr
Galician	Romance (Ibero) (Indo-European)	*gache*	[gat͡ʃe]	[t͡ʃ]	academia.gal/ dicionario
		isca	[iska]	[s]	
German	Germanic (West) (Indo-European)	*husch*	[hʊʃ]	[ʃ]	duden.de own data
		sch	[ʃ]	[ʃ]	
		scht	[ʃt]	[ʃ]	
		ksch	[kʃ]	[ʃ]	
		kscht	[kʃt]	[ʃ]	
Greek	Hellenic (Indo-European)	*ουστ*	[ust]	[s]	own data
		ξουτ	[ksut]	[s]	
Greek Ancient	Hellenic (Indo-European)	*σιττα*	[sitːa]	[s]	Carrington Bolton (1897)
Gusii	Bantu (E) (Niger-Congo)	*shee*	[ʃeː]	[ʃ]	own data
		ksuu	[ksuː]	[s]	
		shiishii	[ʃiːʃiː]	[ʃ]	
		ss	[sː]	[s]	
Hausa	Chadic (Afroasiatic)	*car*	[t͡ʃar]	[t͡ʃ]	Olderogge (1963) own data
		cir	[t͡ʃir]	[t͡ʃ]	
		su	[su]	[s]	

Table 1. cont.

Language	Family	Lexeme(s)	IPA	Sibilant	Source
		shush	[ʃuʃ]	[ʃ] and [ʃ]	
		uss	[us:]	[s]	
Hebrew	Semitic (North-West) (Afroasiatic)	קִשְׁטָה	[kiʃta]	[ʃ]	own data
Hungarian	Finno-Ugric (Uralic)	*hess*	[hɛʃ:]	[ʃ]	mek.oszk.hu
		sicc	[sit͡s:]	[s] and [t͡s]	
		sscc	[s:t͡s]	[s] and [t͡s]	
Indonesian	Malayo-Polynesian (Austronesian)	*hus(s)*	[hus(:)]	[s]	Stevens and Schmidgall - Tellings (2010) own data
		cik	[t͡ʃiʔ]	[t͡ʃ]	
		soh	[soh]	[s]	
Italian	Romance (Italo-Dalmatian) (Indo-European)	*sciò*	[ʃɔ]/[ʃo]	[ʃ]	grandidizionari.it
Japanese	Japonic	し っ し っ (Hiragana) 𠮟𠮟 (Kanji)	[ɕiɕi]	[ɕ]	own data
Karakalpak	Kipchak (Turkic)	*пыш-ай*	[pɯʃ-aj]	[ʃ]	Denisova and Sergeev (2015)
		пыш-хай	[pɯʃ-xaj]	[ʃ]	
Kazakh	Kipchak (Turkic)	*пыш*	[p$^{(h)}$əʃ]	[ʃ]	Denisova and Sergeev (2015) Shorybaev (2014) own data
		брысь	[brəsʲ]	[sʲ]	
		тырш	[təpʃ]	[ʃ]	
		чек	[t͡ʃi̯ek]	[t͡ʃ]	
		қош	[q$^{(h)}$oʃ]	[ʃ]	
		чай	[t͡ʃaj]	[t͡ʃ]	
Konso	Cushitic (Afroasiatic)	*saay*	[sa:j]	[s]	Orkaydo (2013)
		usuk	[usuk]	[s]	
		usss	[us:]	[s]	
		ʃok	[ʃok]	[ʃ]	
Korean	Koreanic	쉬	[ɕi]	[ɕ]	own data
Latvian	Baltic (Indo-European)	*iššš*	[iʃ:]	[ʃ]	Smiltniece (2013)

Table 1. cont.

Language	Family	Lexeme(s)	IPA	Sibilant	Source
		škic	[ʃkit͡s]	[ʃ]	
		tiš	[tiʃ]	[ʃ]	
Lithuanian	Baltic (Indo-European)	*šš*	[ʃː]	[ʃ]	Ambrazas et al. (2006) own data
		št	[ʃt]	[ʃ]	
		tìš	[tɪʃ]	[ʃ]	
		č	[t͡ʃ]	[t͡ʃ]	
		štìš	[ʃtɪʃ]	[ʃ] and [ʃ]	
		šč	[ʃt͡ʃ]	[ʃ] and [t͡ʃ]	
		čìs	[t͡ɕɪs]	[t͡ɕ] and [s]	
		škàc	[ʃkɐt͡s]	[ʃ] and [t͡s]	
		špùl	[ʃpʊɬ]	[ʃ]	
Luganda	Bantu (J) (Niger-Congo)	*shii*	[ʃiː]	[ʃ]	own data
		suu	[suː]	[s]	
Lushai	Tibeto-Burman (Sino-Tibetan)	*sîk*	[sik]	[s]	Lorrain (1940-2005)
Maale	Omotic (South) (Afroasiatic)	*šíkk-…*	[ʃikː]	[ʃ]	Amha (2001; 2013)
Makhuwa	Bantu (P) (Niger-Congo)	*osu*	[osu]	[s]	own data
		suka	[suka]	[s]	
Malay	Malayo-Polynesian (Austronesian)	*syuh*	[ʃuh]	[ʃ]	own data
		syu	[ʃu]	[ʃ]	
		sy	[ʃ(ː)]	[ʃ]	
		ss	[s(ː)]	[s]	
		hasa	[hasa]	[s]	
		asa	[asa]	[s]	
		hus(s)	[hus(ː)]/ [həs(ː)]	[s]	
Marathi	Indo-Iranian (Indo-Aryan) (Indo-European)	शुक्क	[ɕukka]/ [ʃukka]	[ɕ]/[ʃ]	Molesworth (1857-2020)
		शुत्	[ɕut]/[ʃut]	[ɕ]/[ʃ]	
Matses	Panoan	*sh*	[ʃ]/[ʂ]	[ʃ]/[ʂ]	Fleck (2003)
Nahuatl	Eastern Huastecan (Uto-Aztecan)	*sh'*	[ʃʲ]/[ɕ]	[ʃʲ]/[ɕ]	own data

Table 1. cont.

Language	Family	Lexeme(s)	IPA	Sibilant	Source
Ndebele	Bantu (S) (Niger-Congo)	*shu*	[ʃu]	[ʃ]	own data
		shi	[ʃi]	[ʃ]	
Noon	Cangin (Niger-Congo)	*kees*	[kɛːs]	[s]	Soukka (1999)
Norwegian (Bokmål / Ny-norsk)	Germanic (North) (Indo-European)	*husj*	[hʉʂ]	[ʂ]	ordbok.uib.no own data
		hysj	[hʏʂ]	[ʂ]	
		hyss	[hʏsː]	[ʂ]	
Oriya/Odia	Indo-Iranian (Indo-Aryan) (Indo-European)	ହାସ	[hās]	[s]	Praharaj (1931-2021)
Oromo	Cushitic (Afroasiatic)	bis	[bis]/[bɪs]	[s]	own data
		kis	[kis]/[kɪs]	[s]	
		shu	[ʃu]/[ʃʊ]	[ʃ]	
		shut	[ʃut]/[ʃʊt]	[ʃ]	
		shit	[ʃit]/[ʃɪt]	[ʃ]	
		chirr	[t͡ʃir]/[t͡ʃɪr]	[t͡ʃ]	
		che	[t͡ʃe]/[t͡ʃɛ]	[t͡ʃ]	
		machi	[mat͡ʃi]/ [mat͡ʃɪ]	[t͡ʃ]	
Pashto	Indo-Iranian (Iranian) (Indo-European)	چچه	[t͡ʃat͡ʃah]	[t͡ʃ] and [t͡ʃ]	Raverty (1867-2007)
		چخه	[t͡ʃixah]	[t͡ʃ]	
		پشته	[paʃtah]	[ʃ]	
		پَشه	[paʃah]	[ʃ]	
		پِشه	[piʃah]	[ʃ]	
		پَشي	[paʃey]	[ʃ]	
		پِشي	[piʃey]	[ʃ]	
Persian	Indo-Iranian (Iranian) (Indo-European)	کیش	[keʃ]	[ʃ]	Hayyim (1934-2021) Steingass (1892-2021) own data
		غس	[ɢis]	[ʃ]	
		پیش	[piʃ]	[ʃ]	
		چخ	[t͡ʃax]	[t͡ʃ]	
		ش	[ʃ]	[ʃ]	
		س	[s]	[s]	
Polish	Slavonic (West) (Indo-European)	*sio*	[ɕɔ]	[ɕ]	Bańko (2008) Daković (2006)
		sz	[ʂ]	[ʂ]	

Table 1. cont.

Language	Family	Lexeme(s)	IPA	Sibilant	Source
		ks	[ks]	[s]	Wierzbicka (2003) Grochowski (1988) Siatkowska (1976)
		ksz	[kʂ]	[ʂ]	
		kysz	[kiʂ]	[ʂ]	
		psik	[pɕik]	[ɕ]	
		akysz	[akiʂ]	[ʂ]	
		apsik	[apɕik]	[ɕ]	
		ps	[ps]	[s]	
Portuguese	Romance (Ibero) (Indo-European)	*xô*	[ʃo]	[ʃ]	own data
		x	[ʃ]	[ʃ]	
Romanian	Romance (East) (Indo-European)	*huş*	[huʃ]	[ʃ]	dictionarroman.ro (2021)
		uş	[uʃ]	[ʃ]	
Russian	Slavonic (East) (Indo-European)	*ки*	[kʂ]	[ʂ]	Daković (2006)
		кыш	[kiʂ]	[ʂ]	
		куш	[kʲiʂ]	[ʂ]	
sePedi (Northern Sotho)	Bantu (S) (Niger-Congo)	*shu-shu*	[ʃu.ʃu]	[ʃ]	own data
Serbo-Croatian	Slavonic (South) (Indo-European)	*kš*	[kʂ]/[kʃ]	[ʂ]/[ʃ]	Čarkić (2010) Daković (2006)
		iša	[iʂa]/[iʃa]	[ʂ]/[ʃ]	
		iš	[iʂ]/[iʃ]	[ʂ]/[ʃ]	
		šiš	[ʂiʂ]/[ʃiʃ]	[ʂ]/[ʃ]	
		šic	[ʂit͡s]/[ʃit͡s]	[ʂ]/[ʃ] and [t͡s]	
		pis	[pis]	[s]	
		oš	[oʂ]/[oʃ]	[ʂ]/[ʃ]	
Shona	Bantu (S) (Niger-Congo)	*shu-shu*	[ʃu.ʃu]	[ʃ]	own data
Slovak	Slavonic (West) (Indo-European)	*heš*	[ɦeʂ]/[ɦeʃ]	[ʂ]/[ʃ]	slovnik.juls.savba.sk
		kš(š)	[kʂ]/[kʃ]	[ʂ]/[ʃ]	
		kšu	[kʂu]/[kʃu]	[ʂ]/[ʃ]	
		šc	[ʂt͡s]/[ʃt͡s]	[ʂ]/[ʃ] and [t͡s]	
		šic	[ʂit͡s]/[ʃit͡s]	[ʂ]/[ʃ] and [t͡s]	

Table 1. cont.

Language	Family	Lexeme(s)	IPA	Sibilant	Source
Spanish	Romance (Ibero) (Indo-European)	*ch*	[t͡ʃ]	[t͡ʃ]	dle.rae.es Pérez García (1993) own data
		chus-chus	[t͡ʃus.t͡ʃus]	[t͡ʃ] and [s]	
		tuso	[tuso]/ [tusa]	[s]	
		(dial) *sape*	[sape]	[s]	
		oxte/oste	[oste]	[s]	
		oxe/ose	[ose]	[s]	
		ox/os	[os]	[s]	
Suba	Bantu (E) (Niger-Congo)	*shoo*	[ʃoː]	[ʃ]	own data
		shuu	[ʃuː]	[ʃ]	
		see	[seː]	[s]	
		saa	[saː]	[s]	
		shushushu	[ʃuʃuʃu]	[ʃ]	
Swahili	Bantu (G) (Niger-Congo)	*ss*	[sː]	[s]	own data
Swedish	Germanic (North) (Indo-European)	*sjas* / *schas*	[ɧas] ([ʂas]/ [ʃas])	[ɧ] and [s]	svenska.se own data
		sjo	[ɧo] ([ʂo]/ [ʃo])	[ɧ]	
		sch	[ʂ]/[ʃ]	[ʂ]/[ʃ]	
		hysch / *hyssj*	[hʏʂ]/ [hʏʃ]	[ʂ]/[ʃ]	
		hyss	[hʏsː]	[s]	
Tamil	South-Dravidian (Dravidian)	சூ	[t͡ʃu(ː)]/ [t͡ɕu(ː)]	[t͡ʃ]/[t͡ɕ]	Winslow (1862-2010) Ramakrishnan (1992)
Turkish	Oghuz (West) (Turkic)	*kış*	[kɯʃ]	[ʃ]	Johanson (2021) seslisozluk.net sozluk.gov.tr own data
		kışt	[kɯʃt]	[ʃ]	
		pist	[pist]	[s]	
		hoşt	[hɔʃt]	[ʃ]	
Ukrainian	Slavonic (East) (Indo-European)	*кuuu*	[kʲiʃ]	[ʃ]	slovnyk.ua
Urdu	Indo-Iranian (Indo-Aryan) (Indo-European)	چخے	[t͡ʃəxeː]/ [t͡ʃaxeː]	[t͡ʃ]	Fallon (1879-2010)

Table 1. cont.

Language	Family	Lexeme(s)	IPA	Sibilant	Source
					Platts (1884-2015)
Vietnamese	Vietic (Austroasiatic)	*xùy*	[s^{w}i]	[s^{w}]	own data
		xuỵt	[s^{w}iʔ][11]	[s^{w}]	
		suỵt	[ʂ^{w}iʔ]/[ʃ^{w}iʔ]	[ʂ^{w}]/[ʃ^{w}]	
Wolaitta	Omotic (Central) (Afroasiatic)	*šúh*	[ʃuh]	[ʃ]	Amha (2013)
		šúk	[ʃuk]	[ʃ]	
		ʔišík	[ʔiʃik]	[ʃ]	
Wymysorys	Germanic (West) (Indo-European)	*huśa*	[huɕa]	[ɕ]	own data
Xhosa	Bantu (S) (Niger-Congo)	*shu*	[ʃu]	[ʃ]	Andrason and Dlali (2020) Andrason (2022)
		futsek	[futsek]	[s]	
		ts(s)ek	[t͡s(:)ek]	[t͡s]	
		ssek	[s:ek]	[s]	
		s(s)uka	[s(:)uka]	[s]	
Yoruba	Volta-Congo (Niger-Congo)	*ṣuṣ*	[ʃuʃ]	[ʃ]	own data
Zargulla	Omotic (East) (Afroasiatic)	*káis*	[kais]	[s]	Amha (2013)
		ʔiššá	[ʔiʃ:a]	[ʃ]	
		ʔúss	[ʔus:]	[s]	
Zulu	Bantu (S) (Niger-Congo)	*fusek*	[fusek]	[s]	own data

4. Results and discussion

All 79 languages that have been reviewed in my study contain more or less common, entrenched, and/or lexicalized dispersals that are built around voiceless sibilants. In the immense majority of those languages, voiceless-sibilant dispersal calls constitute the most typical and the most salient dispersals – i.e., first-come-to-mind and/or most clearly associated with a directive-to-animal function – or, alternatively, belong to the set of such dispersals. Equally significant is the fact that the majority of the languages, specifically 55 varieties, make use of more

[11] Sometimes, the final <*t*> is realized with a [t̚] that has no audible release.

than one voiceless-sibilant dispersal. This is the most evident in Arusa, Kazakh, Malay, and Persian (6x), Basque, Pashto, Polish, Serbo-Croatian, and Spanish (7x), Ayt Hadiddu and Oromo (8x), Lithuanian (9x), and, especially, Chuvash (14x). In the case of 24 languages that contribute to my database with only one voiceless-sibilant dispersal, such a lexeme again tends to be, by far, the most typical and salient one among all dispersals available (see, e.g., *shoe* in Afrikaans, *su*(*u*) Akan, and *shu*(-*shu*) in Ndebele, sePedi, and Shona).[12] Overall, in the entire language sample, 247 voiceless-sibilant dispersals are attested, which yields an average of 3.1 lexemes per variety.

Notwithstanding the remarkable quantitative visibility of voiceless-sibilant dispersals described above, these types of lexemes are certainly not the unique dispersal forms attested. First, in all languages, (at least some) animals may be chased away by means of expressions that do not contain voiceless sibilants. This is always possible if dispersals are directed to larger species and/or animals that, in particular societies, entertain close relationships with humans. In such cases, dispersals often draw on conatives that are compatible with human referents (see *won* 'away', and *poszedł* or *wynocha* 'go away' in Polish) and derive from original non-conative constructions, e.g., nouns, imperatives, adverbials, small clauses, adpositional phrases (see *ino* lit. 'go' in Maasai and *mka* lit. 'depart' in Xhosa). Since these dispersals are extensions of human-oriented secondary conatives to CACs, they need not exhibit any preference for the presence of voiceless sibilants.[13] Second, in three languages, the dispersal forms that, according to the interviewed speakers, are the most common, entrenched, and lexicalized do not draw on voiceless sibilants. This situation is attested in Catalan (see the more common lexeme *fora*), Korean (see the more common forms 훠이 [hwoi] and 워이 [woi]), and Zulu (see *gowan*, itself a borrowing from the English expression *go on*). As was the case with the former type of exceptions, these CACs are secondary and compatible with human addressees. Third, in exceptional cases, in the set of dispersals that are primary and/or not derived from human-oriented conatives, voiceless-sibilant dispersals are (much) fewer

[12] This lower number of voiceless-sibilant dispersals in some languages often (although not always) seems to be due to the limited scope of sources mentioning CACs in the respective languages.

[13] The distinction between secondary and primary CACs follows the widely recognized distinction between secondary and primary interjections (Ameka 1992a; 2006; Nübling 2001; Stange and Nübling 2014). Secondary CACs derive from lexical classes other than CACs or from analytic sequences that involve non-CAC lexical classes. Although secondary CACs are commonly used in a directive-to-animal function, their relationship with non-CACs is still available, with regard to both their form and meaning. In contrast, primary CACs are used only in a directive-to-animal function. They typically do not derive from non-CAC structures but are employed as CACs from their grammatical birth. Alternatively, their grammaticalization into CACs is so advanced that any link with their non-CACs sources is lost (Andrason and Karani 2021; Andrason 2022).

than dispersals that do not contain these types of consonants. In my sample of languages, this is only attested in Oriya/Odia, where among four primary dispersals, three do not draw on voiceless sibilants: ଧୂଆ [dhūa], ହାତ୍ ହାତ୍ [hāt.hāt], and ହଃହା [hohā].[14]

The contribution of voiceless sibilants to the collected dispersals is also significant from a qualitative perspective. Lexemes in which at least half of the phonetic material is provided by sibilants constitute 36% (89x) of the dispersals in my sample. This is the most patent in 25 tokens (10%) that draw exclusively on sibilants. In 3 of these cases, the lexeme consists of two sibilants, invariably a fricative and an affricate. In 53 dispersals (21%), a sibilant cooccurs with a single non-sibilant consonant or a vowel. Specifically, 14 lexemes (6%) equal a cluster of two consonants, one of them being a voiceless sibilant, while 39 lexemes (16%) draw on a sibilant and a vowel or diphthong. In 9 dispersals (4%), two or three voiceless sibilants cooccur with, respectively, two or three other consonantal or vocalic phones. In further 9 lexemes (4%), the number of voiceless sibilants surpasses the number of other phones (i.e., 2 versus 1 and 4 versus 2). Overall, 17 dispersals (7%) contain two sibilants; one dispersal contains three voiceless sibilants ([ʃuʃuʃu]); and another one contains 4 such consonants ([t͡ʃus.t͡ʃus]).

The extremely prevalent use of voiceless sibilants in dispersals and the considerable contributions of such consonants to the phonetic substance of these types of CACs demonstrates that the correlation between voiceless sibilants and dispersals noticed impressionistically in my previous studies constitutes a robust crosslinguistic tendency. In my opinion, this evident (quantitative and qualitative) trend suggests, in turn, that the voiceless-sibilant foundation of dispersals is not arbitrary. As hypothesized by Bynon (1976: 62), voiceless

[14] It should be noted that I did not analyze all dispersals found in the 79 languages nor did I establish the relationship between voiceless-sibilant dispersals and the other members of the dispersal category – as explained above, I focused on those dispersals that contain voiceless sibilants. This stems from the fact that, with a few exceptions (Bynon 1976; Siatkowska 1976; Amha 2013; Denisova and Sergeev 2015; Andrason and Karani 2021; Andrason 2022), we still lack comprehensive language-specific studies of dispersals that would identify (nearly) all dispersals and distinguish between their primary and secondary types. Therefore, the sets of dispersals that I identified in respective languages during my research are certainly not exhaustive – according to my field data, the number of dispersals in a language should hover around 20 and a large portion of them should be secondary. Furthermore, since the development towards a secondary dispersal is gradual, the number of dispersals in any given language may be even greater, rendering this category, to an extent, "openable". As a result, the typology of dispersals constitutes a much more difficult task than the comparison of 'frog' words conducted by Berlin (1992) and Hays (1994). A language usually has very few nouns with the meaning 'frog' and does not distinguish between their primary and secondary types – a simple dictionary search is thus sufficient. A similar dictionary search is not possible for a dispersal category. (For the limitations of my study see also Section 5.)

sibilants may indeed have some phonetic properties that make them the most suitable for repelling animals.

The non-arbitrariness of voiceless-sibilant dispersals transpires through the profound formal similarity attested by dispersals that are used in genetically and geographically remote languages. These similarities are evident in the lexemes that are made up entirely of voiceless sibilants (see [s(:)] in Arusa, Dholuo, Gusii, Malay, and Persian as well as [ʃ/ɕ/ʂ(:)] in Arusa, German, Polish, Lithuanian, Matses, Nahuatl, and Persian) or contain a voiceless sibilant together with another consonant (e.g., [ks/ʃ/ʂ(:)] in Ayt Hadiddu, Chuvash, German, and Slavonic (i.e., Czech/Polish/Russian/Serbo-Croatian/Slovak)) or a vowel (e.g., [ʃu(:)] in Akan, Dzə, Luganda, Afrikaans, Bantu-S (Ndebele, sePedi, Xhosa), and Oromo), [ʃ/ɕi] in Korean, Ndebele, and Luganda, as well as [ʃ/ɕo/ɔ] in Italian, Portuguese, and Polish). Equally significant are similarities exhibited by more complex dispersals: [HVS] attested in Finnish, Germanic (German, Norwegian, Swedish, and Wymysorys), Indonesian and Malay, Arabic, Oriya/Odia, Slovak, and Ayt Hadiddu; [kVS] attested in most Slavonic languages, Turkish, Oromo, Chuvash, Estonian, Kazakh, Persian, Noon, and Zargulla; [SVk] attested in Arusa, Konso, Maale, Indonesian, Wolaitta, and Xhosa; [SVt] attested in Marathi, Oromo and Vietnamese; [SVR] attested in Oromo and Hausa; [SVS] attested in Serbo-Croatian, Yoruba, Hausa, Hungarian, Lithuanian, and Spanish; and [VST] attested in Estonian, French, and Greek. Perhaps, forms such as *ks(s)t* (Dutch), *kışt* (Turkish), *ξουτ* (Greek), and קִשְּׁטָה (Hebrew), on the one hand, and *pis* (Serbo-Croatian), پیش (Pashto), *bis* in Oromo, *ps(s)t* and *pist* (Danish), *pist* (Turkish), *xipist* (Basque), on the other hand, illustrate the crosslinguistic convergence of dispersals most clearly.

The crosslinguistic convergence of dispersals not only concerns the quantitative and qualitative relevance of voiceless sibilants present in these types of CACs, it is also related to the type of sibilant attested, the overall phonotactic structure, including the position of sibilants, and the properties of vowels and other non-sibilant consonants.

As far as the sibilant type is concerned, out of the 272 sibilants found in the collected dispersals, fricatives predominate (87%). Hushing fricatives ([ʃ], [ʂ], [ɕ], [ʃ̻]) are more common (53%) than hissing fricatives ([s]) (34%). Among all fricative sibilants, the most prevalent are [ʃ] (42%) and [s] (34%). The other types are much less frequent: [ʂ] – 8%, [ɕ] – 4%, and [ʃ̻] – less than 1%. Affricate sibilants are significantly rarer (13%). Similar to fricatives, affricates with a hushing release are more common (10%) than affricates with a hissing release (4%). In further similarity, among all affricate sibilants, the most frequent is [t͡ʃ] (with a [ʃ] release) (8%), the other variant, i.e., [t͡ɕ] (with a [ɕ] release) being much less common (more than 1%). Four types of the coarticulations of sibilants are attested: labio-velar [^{w}], non-velarized labialized (whistled) [^{ɸ}], palatalized [^{j}], pharyngealized [^{ʕ}], and aspirated [^{h}] – all of them infrequent (7x – less than 3%).

Table 2: Sibilant types

sibilant	instances	frequency
[ʃ]	103x+20x[15]	42%
[s]	92x+2x	34%
[t͡ʃ]	22x+1x	8%
[ʂ]	10x+15x	6%
[ts]	10x	4%
[ɕ]	9x+3x	4%
[t͡ɕ]	3x+1x	>1%
[ʃ̺]	2x	<1%

As far as the phonotactics of the 247 voiceless-sibilant dispersals are concerned, the most prevalent structure is mono-syllabicity (59%). The syllables of such dispersals are most commonly closed (107x, i.e., 73% of all monosyllabic lexemes). They may however also be open (39x/27%). Codas consisting of a single element are more common (98x/67%) than complex codas (9x/6%). In most cases, monosyllabic dispersals begin with a consonant (135x/92%). In contrast, onset-less lexemes are only found 11 times (8%). Like codas, onsets built of a single element (sibilant or not) are more common (117x/80%) than complex onsets (18x/12%). In monosyllabic dispersals, sibilants are found equally often in onsets (75x/51% of these lexemes) and codas (76x/52%). Of all such cases, 11x (8%) exhibit sibilants in both codas and onsets. The prevalent syllable structures in monosyllabic lexemes – and in fact in all types of analyzed dispersals – are CVS (32% for monosyllabic tokens and 19% for all tokens), SV (21%/13%), and SVC (18%/11%). They jointly amount to 71%/42%. Other monosyllabic combinations are much less common (29%/17%). Bisyllabicity is attested 55 times (22% of all voiceless-sibilant dispersals). Similar to monosyllabic dispersals, most bisyllabic lexemes end in an open syllable (38x/70% of such lexemes). Bisyllabic lexemes with final closed syllables are much less frequent (16x/30%). In a further similarity to monosyllabic dispersals, most bisyllabic tokens start with a consonant (46x/85%), while onset-less forms are rare (8x/15%). There is only one trisyllabic dispersal (<0.5%), resulting from the triplication of a monosyllable that is open and has a consonantal (sibilant) onset. Additionally, 45 dispersals (18%) are entirely made up of consonants. Such non-

[15] The symbol "+" refers to the cases in which more than one realization or interpretation is possible. Specifically, Table 2 indicates that, in 2 cases, a sibilant can be realized as [s] and [ʃ]. In 18x, [ʃ] is alternatively analyzed as [ʂ], [s̻], [ɕ]; in 1x, [t͡ʃ] is analyzed as [t͡ɕ]; in 15x, [ʂ] is analyzed as [ʃ]; in 3x, [ɕ] is analyzed as [ʃ]; and in 1x, [t͡ɕ] is analyzed as [t͡ʃ].

vocalic lexemes may consist of one phone (9%), two phones (7%), or 3 phones (2%). A longer sequence arises in one instance due to the reduplication or multiplication of a bi-phonic singleton (<0.5%).

Table 3: Syllable structure of voiceless-sibilant dispersals

pattern			instances	frequency
non-vocalic	monosyllabic	S	22x	9%
		CS	11x	4%
		CSC	4x	2%
		SS	3x	1%
		SC	3x	1%
		CSS	1x	0.5%
	plurisyllabic	$CSCS^n$	1x	0.5%
vocalic	mono-syllabic	CVS	46x	19%
		SV	31x	13%
		SVC	26x	11%
		SVS	8x	3%
		VS	8x	3%
		CSV	6x	2%
		CSVC	4x	2%
		SVSC	4x	2%
		VSC	3x	1%
		SCVS	3x	1%
		CCVS	2x	<1%
		CVCS	2x	<1%
		SCV	2x	<1%
		SCVC	1x	0.5%
	bisyllabic	SVCV	13x	5%
		CVSV	11x	4%
		CVSVC	6x	2%
		SVSV	4x	2%
		VSV	4x	2%
		CVCCVS	2x	<1%

Table 3 cont.

pattern			instances	frequency
		CVSCVC	2x	<1%
		VSCV	2x	<1%
		CVCSVC	1x	0.5%
		CVSCV	1x	0.5%
		CVSVS	1x	0.5%
		SCVCV	1x	0.5%
		SVCVC	1x	0.5%
		SVCVSC	1x	0.5%
		SVSSVS	1x	0.5%
		SVSVC	1x	0.5%
		VCSVC	1x	0.5%
		VCVS	1x	0.5%
		VSVC	1x	0.5%
	trisyllabic	CVCVCV	1x	0.5%

As far as the vocalic material is concerned, out of 258 vowels present in the analyzed dispersals, (near-)close vowels are significantly more common (55%) than mid (21%) and (near-)back vowels (25%). Among (near-)close and mid vowels – front and back vowels predominate while central vowels are rare. To be exact, I-type ([i]/[ɪ]/[ʏ]) and U-type ([u]/[ʊ]/[ɯ]) vowels each constitute 27% of all the vowels attested in voiceless-sibilant dispersals, whereas Ɨ-type vowels ([ɨ]/[ʉ]) amount to only 1%. Similarly, E-type ([e]/[ɛ]) and O-type ([o]/[ɔ]/[ɤ]) vowels constitute 10% and 8% respectively, whereas the Ə-type ([ə]/[ɘ]) represent only 2%. Interestingly, the majority of the A-type vowels ([a]/[ɐ]) – the third most common vowel type – are found in bisyllabic lexemes (42x/66%). For the two most common types, U- and I-type vowels, the tendency is opposite and the majority of them appear in monosyllabic lexemes (44x/63% and 43x/62%, respectively). With regard to bisyllabic dispersals, only 15x (27% of these lexemes) attest to vocalic harmony, with the vowel timbre tendencies being similar to those discussed above. That is, (near-)close vowels predominate (9x): U-type – 5x and I-type – 4x. (Near-)open vowels are less common: A-type – 6x. Mid vowels are the least common: O-type – 1x. Out of such harmonious lexemes, 6 arise due to reduplication. In the only trisyllabic lexeme, a U-type vowel is used ([u]).

As far as consonantal material is concerned, out of 194 non-sibilant consonants attested, plosives are by far the most common (72%). Voiceless

Table 4: Vowels found in voiceless-sibilant dispersals[16]

vowel		instances	frequency
U-type	close back	69x+1x	27%
I-type	(near-)close front	69x	27%
A-type	(near-)back	63x+1x	25%
E-type	mid front	25x	10%
O-type	mid back	22x	8%
Ə-type	mid central	5x+2x	2%
Ɨ-type	close central	3x	1%

plosives are much more frequent (66%) than their voiced counterparts (6%). Within plosives, velars constitute 37%, bilabials – 16%, alveolars – 15%, glottals – 4%, and uvulars – 1%. Three voiceless plosives predominate – [k] (35%), [p] (14%), and [t] (13%) – jointly constituting 62% of all consonants. Other types of plosives – i.e., [ʔ], [b], [d], [g], [q], and [ɢ] – are much less frequent. Fricatives are second most common (21%). Guttural [H]-type fricatives prevail (19%), the only other type (labio-dental [f]) amounting to 1%. Among all guttural fricatives, [h] is by far more frequent (14%) than [x], [χ], [ɦ], and [ʁ]. The remaining consonant types are very rare: approximants (only [j]) – 3%, trills, flaps, or taps ([r] and [ɾ]) – 4%; nasals (only [m]) – 1%; and laterals (only [ɬ]) – 0,5%. In replicated vocalic (bisyllabic and trisyllabic) lexemes, no consonant other than the sibilant is attested.[17]

While several aspects of the frequencies provided above need not be universal – my sample of the lexemes is not fully representative and the commonness order of the various patterns and properties, whose occurrences differ minimally, is certainly not fixed – a number of robust tendencies are evident. These tendencies are, in my view, crosslinguistically valid and should hold true when more evidence is accumulated.

In light of these robust tendencies, it is possible to propose the prototype of a voiceless-sibilant dispersal. As usual, the prototype is cumulative and exhibits the following features: the sibilant is a pure hushing phone with no secondary articulation, i.e., [ʃ]; it occupies either a word-initial or word-final position, although CVS is the most common; the lexeme is a monosyllable with an onset

[16] In two cases, two realizations are possible: U/Ə and AE/ƏE.

[17] It should be noted that many [K]-types appear in clusters with sibilants, i.e., as [KS]. I do not consider them affricates, because they are not analyzed so in the respective languages, contrary to [TS] combinations containing the initial element [T], which appear in total 33 times and which are analyzed as affricates.

Table 5: Consonants found in of voiceless-sibilant dispersals

consonant	instances	frequency
[k]	68x	35%
[p]	27x	14%
[h]	27x	14%
[t]	26x	13%
[ʔ]	7x	4%
[r]	7x	4%
[x]	5x	3%
[j]	5x	3%
[b]	4x	2%
[d]	3x	2%
[g]	3x	2%
[χ]	2x	1%
[f]	2x	1%
[m]	2x	1%
[q]	1x	0,5%
[ɢ]	1x	0,5%
[ɦ]	1x	0,5%
[ʁ]	1x	0,5%
[ɾ]	1x	0,5%
[ɬ]	1x	0,5%

and coda consisting of a single element; the vowel is of the I- and U-types, and the non-sibilant consonant is a voiceless plosive, especially one of a [k]-type. The resulting pattern is [kI/Uʃ].[18] In my opinion, the obtainability of such a voiceless-sibilant-dispersal prototype further corroborates the non-arbitrariness of this class of dispersals, argued above in this section.

Apart from supporting the thesis of the non-arbitrariness of some types of CACs proposed by Bynon, the phonetic properties of voiceless-sibilant dispersals identified above also corroborate the validity of the phonetic profile associated with the CAC prototype.

[18] Interestingly, even in non-vocalic lexemes, the [kʃ/ʂ] pattern is the second most common and is only slightly less common than the pattern containing a single fricative sibilant.

With regard to the consonantal nature of CACs, consonants play a significantly more prominent role in voiceless-sibilant dispersals than vowels. To be exact, lexemes in which the number of consonants exceeds the number of vowels, are much more common (171x/69%) than the lexemes exhibiting an opposite relationship (4x/2%).[19] This tendency is visible in both monosyllabic and bisyllabic dispersals. Specifically, 73% of monosyllabic lexemes contain more consonants than vowels.[20] For bisyllabic lexemes, 35% contain more consonants than vowels, while the inverse ratio is only attested in 7%.[21] Probably, the most evident exponent of the consonantal nature of voiceless-sibilant dispersals is the common presence of non-vocalic lexemes (18% of the total), i.e., lexemes that consist of sole consonants.

With regard to extra-systematicity, sounds that are not included in the IPA alphabet, in particular whistles and kissing sounds, are unattested in voiceless-sibilant dispersals. This likely stems from the fact that whistles and kissing sounds generally do not combine with proper speech phones in CACs (see Andrason and Karani 2021). Sounds that are foreign to the language in which a given dispersal is used but are included in the IPA alphabet and belong to the repertoire of human phones are also absent in voiceless-sibilant dispersal. This may again be explained with the help of two other facts. First, sibilants – especially the alveolar one, but also hushing ones, either fricatives or affricates – are relatively common cross-linguistically (Ladefoged and Maddieson 1996). Indeed, in all languages included in this research, the alveolar sibilant forms part of a standard phonetic and/or phonological inventory. If a language has no [ʃ] sibilant – the most prototypical sibilant used in dispersals according to my data – it resorts to other available sibilants that are acoustically similar, i.e., hushing phones [ɕ] and [ʂ] (e.g., Polish), including affricates with a hushing release, i.e., [tʃ] (e.g., Spanish), or, always possible, a hissing phone, [s] (e.g., Dholuo). As a result, there is no inherent need for an IPA phone that would be absent in a particular language. Second, the most characteristic extra-systematic IPA sounds found in CACs across languages are clicks (Andrason and Karani 2021). However, since clicks rarely combine with other phones in non-click languages, it is hardly surprising that no clicks are attested in voiceless-sibilant dispersals.[22] While extra-systematic vowels and consonants are absent, voiceless-sibilant dispersals exhibit a considerable extent of phonotactic extra-systematicity. That is, they tolerate phonotactic structures that are otherwise rare or absent in a given language. This includes complex consonant clusters (see, for instance, [ts] in

[19] For nearly a third of the lexemes (29%), the contribution of consonants and vowels is equal.
[20] In the remaining 27% of monosyllabic lexemes, the number of consonants and vowels is equal.
[21] In the remaining 58% of bisyllabic lexemes, the number of consonants and vowels is equal.
[22] This, however, does not mean that clicks are never used in dispersals, see, e.g., [ǁ] in Persian.

Estonian or [bs(ː)d] in Danish, as well as a word-final consonant [k] in Xhosa and Zulu), non-vocalic lexemes (found in 33% of languages), and extra-long consonants (see next paragraph). In the respective languages, these three features are disallowed or dispreferred in the lexical classes other than interjections, ideophones, and conatives.[23]

With regard to the phonotactic, suprasegmental, and phonation-related properties associated with the prototype of dispersals, the following should be noted: First, voiceless-sibilant dispersals attest to phonetic shortness and monosyllabicity. That is, lexemes consisting of less than 2 syllables clearly predominate (188x/77%). In contrast, bisyllabic lexemes are much rarer (55x/22%), while longer structures are exceptional (1x/<0.5%) and only arise due to the replication of monosyllabic (vocalic or non-vocalic) singletons. Replications themselves, although attested, are uncommon (8x/3%). A more frequent type of extension is prolongation, especially the prolongation of sibilants. To be exact, 42x dispersals (17%) are pronounced with a long or extra-long sibilant and, in all the other languages the native speakers of which I consulted, without being the most entrenched, this (extra-)long pronunciation of sibilants is always grammatical in voiceless-sibilant dispersals. Additionally, voiceless-sibilant dispersals widely attest to optional repetitions. That is, in all languages included in my survey, the collected voiceless-sibilant dispersals may – and very often are – repeated, usually in sequences of two or three. Second, voiceless-sibilant dispersals comply with the modulations associated with the prototype of dispersals, especially articulatory speed or short rate of production and intense phonation or a raised aggressive voice. The articulatory speed is compatible with the short phonetic substance, while the intensity is compatible with the lengthening of the sibilant – both phenomena noticed above. Overall, voiceless-sibilant dispersals tend to be pronounced with a powerful egression of airstream, which may end abruptly or persist for a time.[24] This phonation-related behavior or tense and aggressive realization is another extra-systematic property characteristic of voiceless-sibilant dispersals.

To conclude, the prototype of CACs, in general, and that of dispersals, in particular – which have been recently formulated in scholarship (Andrason and Karani 2021; Andrason 2022; Heine 2023) – may be regarded as accurate. Voiceless-sibilant dispersals meet nearly all prototypical features. The only divergence, i.e., the absence of extra-systematic sounds, stems from the properties of such extra-systematic sounds themselves and the nature of the most salient component of the dispersals analyzed in the present study, i.e.,

[23] For an additional phonation/intonation-related extra-systematic property, see next paragraph.
[24] This is often reflected in the descriptions found in grammars and dictionaries, where the sibilants used in dispersals are described as emphatic or strong (see dle.rae.es for Spanish and svenska.se for Swedish examples).

sibilants. Therefore, the observed incompatibility with extra-systematic sounds likely pertains to only voiceless-sibilant dispersals – it does not undermine the ability of other dispersals and CACs to host such sounds, nor does it compromise the validity of the CAC and dispersal prototype.

5. Conclusion

The present study demonstrated that languages tend to contain dispersals that are built around voiceless sibilants. This tendency is both quantitative (i.e., voiceless-sibilant dispersals are common across languages and in a single language) and qualitative (i.e., sibilants contribute very significantly to the phonetic substance of such dispersals). This fact, together with a range of formal similarities exhibited by voiceless-sibilant dispersals in the languages of the world which led to the formulation of their prototype – encapsulated by the pattern [voiceless plosive + (near-)close vowel + sibilant], specifically, [kI/Uʃ] – suggests that the presence of voiceless sibilants in dispersals is not arbitrary. As hypothesized by Bynon (1976), voiceless sibilants may produce some acoustic effect that render them suitable for repelling bothersome animals. Furthermore, voiceless-sibilant dispersals tend to comply with the general phonetic profile associated with the prototype of CACs and dispersals, thus corroborating the validity of this prototype.

While I have responded to the research question formulated at the beginning of this article, I have certainly not exhausted all issues relates to voiceless-sibilant dispersals. First, in this article, I have not studied the semantic differences between the various voiceless-sibilant dispersals. Given that CACs tend to be specialized with regard to their meaning, in particular, being often used with determined species or types of animals (Andrason and Karani 2021: 34), voiceless-sibilant dispersals may show certain species-related preferences. Second, I have not analyzed similarities exhibited by voiceless-sibilant dispersals within a single language family or across an adjacent geographic area. Given that CACs can be both inherited and borrowed (ibid. 35), it is possible that they reveal genetic and areal tendencies. Third, I have not determined the role which voiceless-sibilant dispersals play within the entire category of dispersals. Given the typological commonness of voiceless-sibilant dispersals described in the present article, this class of dispersals may be the most prevalent among all dispersal types. However, as the other types of dispersals have not been systematically examined in this research, such a generalization remains more or less impressionistic (see footnote 14 below). Fourth, and related to the above point, I have not studied specifically the presence and/or prevalence of sibilant dispersals in languages that lack sibilants in their standard sound inventory (e.g., Dinka and Lango in Africa or several native Australian languages). In light of the

abovementioned limitations, my future research will focus on the study of the semantic and genetic/areal properties of voiceless-sibilant dispersals, their relationship to the remaining dispersal types, and their potential use (or absence) in non-sibilant languages.

Additionally, while the crosslinguistic convergence regarding the phonetics of dispersals – i.e., the use of voiceless sibilants and their cooccurrence with voiceless plosives and (near-)close vowels – is unmistakable and the non-arbitrariness of these dispersals, in my view, is evident, the exact motivation of the presence of voiceless sibilants in dispersals remains an open question. Are sibilants used in dispersals because animals dislike such phones? Remarkably, onomatopoeias imitating the sound made by snakes draw on sibilants in many languages (see, e.g., Poyatos 1993; 2002; Bańko 2008), which suggests that sibilants may indeed be associated by (some) animals with danger.[25] Are sibilants used in dispersals because they trigger some negative connotations for humans, which then leads to their subjective association with potential repelling effects on other species? For example, Dionysius of Halicarnassus wrote already in the 1st c. BCE that "σ [i.e., *s*] [was] an unattractive, disagreeable letter, positively offensive when used to excess, [because representing] a sound more suited to a brute beast than to a rational being" (Roberts 1910: 146). Similarly, Jakob Boehme suggested in the 17th c. that "the sibilant, hissing sound […] "sch" or "f" was […] associated with the devil" (Coudert 1998: 97).[26] Or are sibilants used in dispersals because they best lend themselves to the exploitation of other phonation-related features which are the true elements repelling animals. As noted by Ladefoged and Maddieson, "sibilants […] have a greater intensity – they are louder – than the other two voiceless fricatives" (1996: 57), "have more acoustic energy" (ibid. 168), and therefore "stand out from other sounds even in non-speech contexts" (ibid.).

The three motivations for the use of voiceless sibilants in CACs suggested above (i.e., biological, psycholinguistic, and acoustic/articulatory/auditory) need not be viewed as mutually exclusive. In my opinion, they are closely related and jointly explain why humans resort to these particular phones when they want to chase away other species.

[25] It should however be noted that in snake onomatopoeias, the typical sibilant is the hissing [s] contrary to dispersals, where a hushing sibilant prevails.

[26] The idea of relating voiceless-sibilant dispersals to the danger associated with snakes as well as the above quotation from Dionysius of Halicarnassus have been suggested to me by Prof. Mirosław Bańko. The reference to Coudert's work was suggested to me by an anonymous reviewer.

References:

Abdulla, A.B., and Z. Nasir Talib 2009. The meanings of interjections in English and Arabic. *Journal of the College of Arts University of Basrah* 50: 89-107.

Alcover, A.M., and F. de Borja Moll 1969. *Diccionari català-valenciá-balear*. Palma de Mallorca: Editorial Moll.

Ambrazas, V., E. Geniušienė, A. Girdenis, N. Sližienė, D. Tekorienė, A. Valeckienė and E. Valiulytė 2006. *Lithuanian grammar*. Vilnius: Baltos Lankos.

Ameka, F. 1992a. Interjections: The universal yet neglected part of speech. *Journal of Pragmatics* 18(2/3): 101-118.

Ameka, F. 1992b. The meaning of phatic and conative interjections. *Journal of Pragmatics* 18(2/3): 245-271.

Ameka, F. 2006. Interjections. In K. Brown (ed.), *Encyclopedia of language and linguistics*, 743-746. Oxford: Elsevier.

Amha, A. 2001. *The Maale language*. Leiden: CNWS.

Amha, A. 2013. Directives to humans and to domestic animals: The imperative and some interjections in Zargulla. In M-C. Simeone-Senelle and M. Vanhove (eds.), *Proceedings of the 5th international conference on Cushitic and Omotic languages, Paris, 16-18 April 2008*, 211-229. Köln: Rüdiger Köppe.

Andrason, A. 2022. Conative animal calls in Xhosa: Testing the prototype. *Africana Linguistica* 28. 25-54.

Andrason, A., and M. Dlali 2020. The (crucial yet neglected) category of interjections in Xhosa. *STUF – Language Typology and Universals* 73(2): 159-217.

Andrason, A., and M. Karani 2021. Conative calls to animals: From Arusa Maasai to a cross-linguistic prototype. *Łódź Papers in Pragmatics* 37(1): 3-40.

Bańko, M. 2008. *Współczesny polski onomatopeikon. Ikoniczność w języku*. Warszawa: PWN.

Berlin, B. 1992. *Ethnobiological classification: Principles of categorization of plants and animals in traditional societies*. Princeton: Princeton University Press.

Bolton, C. 1897. The language used in talking to domestic animals. *The American Anthropologist* 10: 65-90.

Bynon, J. 1976. Domestic animal calling in a Berber tribe. In W. McCormack and S. Wurm (eds.), *Language and man: Anthropological issues*, 39-66. The Hague: Mouton.

Coudert, A. 1998. *The impact of the Kabbalah in the seventeenth century: The life and thought of Francis Mercury van Helmont*. Leiden: Brill.

Čarkić, M. 2010. *On poetic language*. Belgrade: Institute for the Serbian Language.

Daković, S. 2006. *Interiekcje w języku polskim, serbskim, chorwackim i rosyjskim. Opis i konfrontacja*. Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.

Denisova T., and L. Sergeev 2015. О некоторых особенностях верхового диалекта в говорах чувашской диаспоры Ульяновской области. *Вестник ЧГПУ им. И. Я. Яковлева* 85(1): 32-37.

Dixon, R.M.W. 2010. *Basic linguistic theory. Methodology. Vol 1*. Oxford: Oxford University Press.

Dryer, M. 2006. Descriptive theories, explanatory theories, and Basic Linguistic Theory. In F. Ameka, A. Dench and N. Evans (eds.), *Catching Language. The Standing Challenge of Grammar Writing*, 207–234. Berlin: Mouton de Gruyter.

Evans, V., and M. Green 2006. *Cognitive linguistics: An introduction*. Edinburgh: Edinburgh University Press.

Fallon, S. 1879-2010. *A new Hindustani-English dictionary, with illustrations from Hindustani literature and folklore*. Banaras, London: Medical Hall Press. https://dsal.uchicago.edu/dictionaries/fallon/

Fleck, D. 2003. *A grammar of Matses*. Houston: Rice University dissertation.

Fried, M., and J-O. Östman 2004. Construction Grammar: A thumbnail sketch. In M. Fried and J-O. Östman (eds.), *Construction Grammar in a cross-language perspective*, 11-86. Amsterdam: John Benjamins.

Goldberg, A. 2003. Constructions: A new theoretical approach to language. *Trends in Cognitive Science* 7(5): 219-224.

Grochowski, M. 1988. Wprowadzenie do analizy syntaktycznej wykrzykników. *Polonica* 13: 85-100.

Hays, T. 1994. Sound symbolism, onomatopoeia, and New Guinea frog names. *Journal of Linguistic Anthropology* 4: 153-174.

Hayyim, S. 1934-2021. *New Persian-English dictionary, complete and modern, designed to give the English meanings of over 50,000 words, terms, idioms, and proverbs in the Persian language, as well as the transliteration of the words in English characters. Together with a sufficient treatment of all the grammatical features of the Persian Language*. Teheran: Librairie-imprimerie Béroukhim. https://dsal.uchicago.edu/dictionaries/hayyim/

Heine, B. 2023. *The grammar of interactives*. Oxford: Oxford University Press.

Hoffmann, T., and G. Trousdale 2013. Construction Grammar: Introduction. In Th. Hoffmann and G. Trousdale (eds.), *The Oxford handbook of Construction Grammar*, 1-12. Oxford: Oxford University Press.

Janda, L. 2015. Cognitive linguistics in the year 2015. *Cognitive Semantics* 1: 131-154.

Johanson, L. 2021. *Turkic*. Cambridge: Cambridge University Press.

Ladefoged, P., and I. Maddieson 1996. *The sounds of the world's languages*. Cambridge: Wiley-Blackwell.

Lorrain, H. 1940-2005. *Dictionary of the Lushai language*. Calcutta: Asiatic Society. https://dsal.uchicago.edu/dictionaries/lorrain/

Molesworth, T. 1857-2020. *A dictionary Marathi and English*. Bombay: The Bombay Education Society's Press. https://dsal.uchicago.edu/dictionaries/molesworth/

Nübling, D. 2001. Von oh mein Jesus! zu oje! – Der Interjektionalisierungspfad von der sekundären zur primären Interjektion. *Deutsche Sprache* 29(1): 20-45.

Olderogge, M. 1963. *Kamus na Hausa-Rashanci*. Moscow: Ethnographic Institute of the Academy of Science of USSR.

Orkaydo, O.O. 2013. *A grammar of Konso*. Utrecht: Lot.
Pérez García, E. 1993. *Sintiendo riópar diccionario*. Servicio de Publicaciones. Diputación Provincial de Albacete.
Platts, J. 1884-2015. *A dictionary of Urdu, Classical Hindi, and English*. London: W. H. Allen & Co. https://dsal.uchicago.edu/dictionaries/platts/
Poyatos, F. 1993. *Paralanguage: A linguistic and interdisciplinary approach to interactive speech and sounds*. Amsterdam: John Benjamins.
Poyatos, F. 2002. *Nonverbal communication across disciplines. Volume II: Paralanguage, kinesics, silence, personal and environmental interaction*. Amsterdam: John Benjamins.
Praharaj, G. 1931-2021. *Purnnachandra Odia Bhashakosha*. Cuttack: Utkal Sahitya Press. https://dsal.uchicago.edu/dictionaries/praharaj/
Ramakrishnan, S. 1992. *Kriyāvi̲n ta̲rkālat Tami̲l akarāti: Tami̲l-Tami̲l-āṅkilam*. Madras: Kriyā. https://dsal.uchicago.edu/dictionaries/crea/
Raverty, H. 1867-2007. *A dictionary of the Pukhto, Pushto, or language of the Afghans: With remarks on the originality of the language, and its affinity to other oriental tongues*. London: Williams and Norgate, 1867. https://dsal.uchicago.edu/dictionaries/raverty/
Roberts, R. 1910. *Dionysius of Halicarnassus. On literary composition being the Greek text of the De Compositione Verborum*. London: Macmillan and Co.
Shorybaev, S. 2014. Қазақ тіл біліміның мәселелері. Almaty: Abzal-ai.
Siatkowska, E. 1976. *Zachodniosłowiańskie zawołania na zwierzęta: stan obecny, funkcje historyczne, stosunek do systemu językowego*. Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.
Smiltniece, G. 2013. Izsauksmes vārds (interjekcija). In D. Nītiņa and J. Grigorjevs (eds.), *Latviešu valodas gramatika*, 679-691. Rīga: Zvaigzne.
Soukka, M. 1999. *A descriptive grammar of Noon, a Cangin language of Senegal*. London: University of London dissertation.
Stange, U., and D. Nübling 2014. Interjections. In C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill and S. Tessendorf (eds.), *Body – language – communication*, 1982-1989. Berlin: Mouton de Gruyter.
Steingass, F. 1892-2021. *A comprehensive Persian-English dictionary, including the Arabic words and phrases to be met with in Persian literature*. London: Routledge & K. Paul. https://dsal.uchicago.edu/dictionaries/steingass/
Stevens, A., and A.E. Schmidgall-Tellings 2010. *A comprehensive Indonesian-English dictionary*. Athens: Ohio University Press.
Taylor, J. 1995. *Linguistic categorization: Prototypes in linguistic theory*. Oxford: Oxford University Press.
Wierzbicka, A. 2003. *Cross-cultural pragmatics: The semantics of human interaction*. Berlin: De Gruyter.
Winslow, M. 1862-2010. *A comprehensive Tamil and English dictionary of High and Low Tamil*. Madras: P.R. Hunt. https://dsal.uchicago.edu/dictionaries/winslow/