

## Vocalism as a discourse-functional category (Jack W. DuBois)

Vocalism represents an important, yet surprisingly little understood, category of phenomena in the realization of spoken language. Although individual instances of vocalism have been noticed recurrently in the literature, and several have attracted considerable attention in their own right, it must be said that methodological treatment and representation in transcriptions has been haphazard; and most problematically, appreciation of their categorical coherence and hence larger theoretical significance has been limited in scope. Typically, each new vocalism encountered is seen as a unique phenomenon unto itself. Conversely, there is little awareness of the commonalities of form and function that would unite vocalisms in a natural class. It remains a challenge to recognize a unified category which is capable of elucidating the general significance of vocalisms in the realization and interpretation of spoken language in use.

This handout seeks to describe in general terms the phenomenon of vocalism; to clarify its coherence as a discourse-functional category; and to explore the significance of vocalism for understanding spoken language in its own terms, as speaking that is embedded in the embodied actions of speakers engaged in interaction, and which is integrally designed to fit with their capabilities, needs, and goals. While a full exposition of a category as broad, diverse, and open-ended as vocalism is beyond the scope of this handout, and its individual members are so numerous that we will be able to mention little more than mention them, I will present enough examples to give an idea of the scope and significance of the class. The emphasis is on the coherence of the category as a whole, and its unified role in the functional realization and organization of language in use.

Vocalisms can be defined as sounds which are (1) produced with the same apparatus that produces speech (i.e. the human vocal tract), but (2) which are not a part of speech per se. Because vocalisms are made where language is made, they alternate with language in the voicestream of a given individual speaker. As a consequence, discourse participants seem to pay very close attention to vocalisms. Vocalisms are interpreted as sharing the same timestream that organizes speech and other aspects of spoken interaction. This gives them a special significance for sequential organization of interaction. The rich diversity of vocalism categories, including laughter, inhalation, exhalation, clicks, sniffs, coughs, and so on, allows them to carry important information about the state of the person who is producing the vocalism. For all these reasons it is important to have an effective way to represent vocalisms so as to capture their structural and functional roles in language use.

For example, some vocalisms act as an important cue to turn-taking, as when a participant makes an audible in-breath just before starting to speak. Because this practice is so regular, audible in-breath becomes a way for speakers to actively signal that they are seeking to begin a turn. While breathing might seem to be motivated by simple necessity, it is necessary to distinguish audible from inaudible breathing, because of the different interactional significance of each. The fact is that, barring respiratory problems or conditions of physical exertion, breathing is often effectively inaudible (or at least very quiet). Conversational participants are able to get most of their breathing done without making significant audible noise—except when they are about to take a turn, when the in-breath is often done in a way that makes it audible. This gives special meaning to the audible in-breath when it is performed.

The fact that the prospective speaker is *choosing* to make his or her in-breath audible often signals a cue for impending entry into speaking. Breathing has other roles as well (as a reliable indicator of IU boundaries, for example). This is just one example; other vocalisms may also turn out to have functions which go beyond their obviously utilitarian ones (think of the potential for the significant cough).

## Frequency

The following table presents the most important vocalisms by frequency, based on those used in transcribing the Santa Barbara Corpus of Spoken American English. These conventions represent most of the vocalisms that a transcriber is likely to encounter. Table 1 lists vocalisms by overall frequency. The frequencies cited below give the number of occurrences in Parts 1&2 of the SBCSAE (roughly 100,000 words). (Note that this table, as well as Table 2 below, represent vocalisms considered as discrete vocal events, written as separate “words”. As such they do not include vocal modifications of words, such as a glottalized cut-off of a truncated word, or laughing during a word.

**Table 1: Top ten vocalisms in SBCSAE, by frequency**

<b>n</b>	<b>Symbol</b>	<b>vocalism</b>
5474	@	Laugh (all forms)
4934	(H)	Audible inhalation
549	(TSK)	Alveolar click
474	%	Glottal/laryngeal constriction
450	(Hx)	Audible exhalation
148	(SNIFF)	Sniff
156	(AHEM)	Hemming and hawing, or throat clearing (THROAT)
115	(COUGH)	Cough
54	(LAUGH)	Generalized laughter (e.g. of a crowd)
52	(GULP)	Swallowing sound

## Type

It's useful to classify vocalisms according to type, in order to get a more general picture of how they function. One such classification distinguishes the following categories, according to the manner or location in which the sound is made, or the type of iconic or indexical sign involved. Table 2 groups the various vocalisms categories into these broad types.

**Table 2: Vocalisms by type**

<b>Laugh</b>	<b>Air</b>	<b>Larynx</b>	<b>Click</b>	<b>Nose</b>	<b>Onomot.</b>	<b>Expressive</b>
@	(H)	%	(TSK)	(SNIFF)	(POW)	(SCREAM)
@word	(Hx)	(AHEM)	(KISS)	(SNORT)	(BBBB)	(YELL)
(LAUGH)	(SIGH)	(COUGH)	(TCHK)	(SNEEZE)	(BUZZ)	(HOWL)
	(GASP)	(GULP)		(SNORE)		(GROWL)
	(BLOW)	(YAWN)		(SNUFFLE)		(WHISTLE)
		(GROAN)				
		(CHOKE)				
		(SOB)				
		(GRUNT)				
		(BURP)				
		(BELCH)				

(1) ((*A Book about Death* SBC005: 828.73-835.67))

1 PAMELA: My favorite word when I was twelve,  
2 (1.5)  
3 was par[adox].  
4 DARRYL: [(YAWN)]:: <YAWN> Why </YAWN>.

(2) ((*Doesn't Work in This Household* SBC019: 687.625-701.384))

1 JAN; (YAWN):: \_<3.0>  
2 BRETT; You could practically do that with a quarter-sized violin.  
3 (0.6)  
4 Fit it in your hand.  
5 JAN; (SNIFF)  
6 (0.8)  
7 (SIGH)  
8 (TSK) (H) Bob came by this afternoo:n?  
9 FRANK; Bob who.

(3) ((*Risk* SBC024: 738.960-752.380))

1 DAN; So it's just you and me babe.  
2 (1.3)  
3 JENNIFER; What's new.  
4 (1.2)  
5 DAN; @@@@ @  
6 JENNIFER; (KISS)  
7 DAN; (H):[:]  
8 JENNIFER; [(KISS)]  
8 (0.7)  
9 (KISS)  
10 (1.1)  
11 DAN; (KISS)  
12 (0.6)  
13 (TSK) (H)  
14 (0.8)  
15 Oh,  
16 let's see.

(4) ((*Cuz* SBC006: 1560.33-1567.93))

1 ALINA; He's real wishy-washy.  
2 (0.2)  
3 (H) (SNIFF) He's too: uh,  
4 .. namby pamby.  
5 (0.4)  
6 You need somebody to come in and go,  
7 .. (POW).  
8 .. (POW).  
9 .. (POW).

10 (0.4)  
11 LENORE; Yeah:.

## The Vocalism Event: Iconic Representations

It is useful to represent in a quasi-iconic manner the sound of vocalism events as they are experienced by participants. Each distinct audible event token in the recording should be represented by its own distinct notational token. Thus instead of writing a transcriber's comment to the reader, of the form “((Jennifer yawns three times))”, it is preferable to represent the event in a way that more iconically reflects the production of sound by the speaker, as it would be heard by participants, with a separate symbol for each yawning event:

(5) ((*Risk* SBC024: 358.822-367.586))

1 JENNIFER; Every time you take over a country you get cards.  
2 (2.3)  
3 (YAWN) (YAWN) (2.0) (YAWN)

In general, the idea is that each vocalism notation token, such as (YAWN), should represent the actual concrete sound event itself, rather than presenting merely an abstract description of the sound, or of the action that produces the sound. For this reason, a notation that captures a more iconic evocation or “image” of the sound event is to be preferred. This is one reason why we try to use words which are themselves more or less iconic or sound-symbolic (yawn, sniff), and we avoid inflecting them grammatically (i.e., avoiding “yawning” or “yawns”). This representational practice follows the usual practice of cartoonists, who have developed a very effective set of informal notational conventions for evoking the audible form of vocalism events, which are important to represent because of what they tell you about the speaker in the moment.

## Vocalisms vs. Speech

Vocalisms should be clearly differentiated from speech in a transcription (again, cartoonists follow this practice rigorously). This is because participants hear vocalisms as different from lexical words, and respond to them in very different from lexical words, and respond to them in very different ways. In the following example, there is a lexical word representing a (fake) laugh, written as “ha”:

(6) ((*Appease the Monster* SBC013: 1064.05-1067.38))

1 KENDRA; you won't make half of what I make.  
2 Ha ha.  
3 (H)  
4 KEN; .. What:t?

The fake ha-ha laugh cannot be considered a true vocalism. As such it contrasts with a real laugh, which is written as a vocalism using the @ sign.

As another example, sometimes a speaker will speak the word “cough” in a sort of pretense of coughing, but in a way which clearly involves the lexical word “cough” rather than the sound of an actual cough. The lexical word “cough” is written in a normal case rather than in all capitals as for the actual sound of a cough (COUGH):

(7) ((*Appease the Monster* SBC013: 1056.23-1063.33))

1 KENDRA; (H) I'll be making well over seventy,  
2 I'll be making a lot more than hi:[m].

3	JENN;	[ʊ]l <sub>2</sub> .n <sub>2</sub> l.
4	MARCI;	[ <sub>2</sub> Cough <sub>2</sub> ].
5		choke,
6		sputter,
7		Well over seventy:?

One piece of evidence that an act like kissing matters to participants is that participants may occasionally refer to the act overtly. Sometimes the word kiss will occur both as an action, and as a participant's description of the action. In the following example, the lexical word "kissing" written in normal case appears shortly after an actual kiss, represented as a vocalism in all caps:

(8) ((*Risk* SBC024: 1070.483-1097.998))

1	JENNIFER;	(H) (SIGH)
2		Now,
3		if I can just get to the next --
4		(5.1)
5	DAN;	((MOUSE-CLICKS))
6	JENNIFER;	(KISS)
7		(1.2)
8		(KISS)
9		(1.1)
10	DAN;	((MOUSE-CLICKS))
11		Thank you very much.
12		(0.9)
13	JENNIFER;	For what.
14		(0.2)
15	DAN;	I just took over Iceland.
16		(0.3)
17	JENNIFER;	(TSK) Oh.,
18		fuck.
19		(1.0)
20	JENNIFER;	(H) Right while I was <b>kissing</b> you,
21		You: traitor.

Since the set of possible vocalisms is more or less open-ended, and even the more frequent vocalisms are quite varied, it becomes a practical necessity to distinguish them mnemonically using words, rather than giving each a unique symbol. This also allows transcribers to invent new vocalism notations as needed. But in choosing these notations, it is preferable to draw on words from very basic vocabulary with are short, monosyllabic, onomatopoetic, and uninflected. Such words are easiest to "hear" as if they represent the actual audible event the stand for.

Vocalism notations should be used consistently. Once it is decided to represent the sound of swallowing by (GULP), for example, this notation should be used throughout one's transcriptions.

One outstanding representational issue concerns the choice of vocalism for the transcription of languages other than English. Depending on the publications conditions and the target audience for one's transcriptions, it may be necessary to develop vocalism notations appropriate to the language you are transcribing, using the appropriate words in the respective language. For example, you could use the word in the target language that onomatopoetically captures the yawn, cough, sniff, or other vocalism category.

