

Response cries inviting an alignment: Finnish *huh huh*

Samu Pehkonen

Abstract

This study looks at the use of the Finnish response cry *huh huh* as a device to invite an alignment and to assess joint physical experience. Response cries establish a speaker's orientation to the world but they also give cues for the recipient on how to respond to the responsive actions appropriately. While the meaning of *huh huh* is unproblematic for the recipients, both speakers and recipients orient to the vagueness of *huh huh* through their turn-design and turn-formulation. Of particular interest are those rare *huh huh* turns that are repeated by the receiver with prosodic upgrading or downgrading compared to the initial ones. Initiating a sequence with a response cry and repeating it in the second position prove to be effective devices for the participants to maintain social solidarity in physiologically or emotionally dense moments. The data is in Finnish with English translations.

Keywords: response cries; repetition; responsive action; conversation analysis; phonetics

Introduction

This study looks at the Finnish exhaled response cry *huh huh* (produced with close back vowel [u] and thus not to be confused with the orthographically similar universal repair-initial *huh?*, Dingemanse, Torreira, & Enfield, 2013) appearing in transition from terminating a joint physical activity to assessing that activity. The article proposes a recognizable interactional practice where participants externalize their physiological state of “being out of breath” in some sense through a (non)lexical item *huh huh* rather than a more elaborated assessment. For this argument, Goffmanian response cries (Goffman, 1978) are re-specified in three ways: First, while *huh huhs* are triggered by some observable events in the world, they are not so much flooding responses as carefully designed in terms of their temporal arrangement. Second, while *huh huh* has a rather conventionalized lexical form expressing a variety of emotions such as relief, fatigue or surprise (VISK §856), the remaining vagueness in its meaning is solved by the participants within the ecology of the ongoing activity and through phonetic variation. Third, while response cries make speakers’ “presumed inward state” (Goffman, 1978, p. 794) available to co-participants, they also serve to initiate sequences of talk where an aligning response to the initial response cry becomes relevant.

Huh huh can appear stand-alone as a turn or it can be followed by an elaborative turn construction unit (TCU) that further explicates the speaker’s stance. In either case, an affect-related response from a recipient becomes relevant. Of particular analytical interest are those sequentially coupled response cries that are repeated by the receiver in the second position, thus forming a two-turn minimal sequence: such as,

((triggering event))

A: *huh huh*.

B: *huh huh*.

Let us straightforwardly assume that A's *huh huh* is a response to the triggering event equally available to both A and B. In this case then what about B's *huh huh*: Is it an independent response to the triggering event, which was only externalized second to A's turn? Alternatively, are we inclined to see it—also or primarily—as an aligning and affiliating response to A's stance?

Sequences like this where the first response cry is repeated in the next turn are intrinsically linked to B's experiential access and entitlement to A's stance (Heritage, 2011). Because *huh huh*, like other response cries and interjections, has partially conventionalized meaning, the participants need to solve the practical problem of designing their turns and responding to those turns in a way that maintains intersubjectivity and social solidarity.

The proposition starts from situating the sequence in the ongoing social action and paying attention to the suprasegmental elements of response cries. Response cries and tokens express emotions extensively through prosody (e.g., Heritage, 2011, p. 176; Wharton, 2003). Wilson and Wharton (2006, p. 1564) propose that prosody can convey this information both unintentionally and intentionally, the latter option further covertly or overtly. This suggests that response cries can accommodate both conventionalized meaning (how something feels) and the speaker's stance (what should be done about it) and that the local context provides the receiver with sufficient information on what type of responsive action is preferred. The article discusses the recognizability of the practice for the participants by combining conversation analysis with phonetic analysis concerning the realization of *huh huh* as a conventionalized and stylized lexical item. As a result, we should be able to see what sort of interactional work *huh huh* is doing.

The article starts with an analysis of *huh huh* in sequence openings and moves towards rare cases of repeated (or slightly modified) *huh huh* pairs. The next section elaborates the base for the argument by introducing the ways in which previous studies in phonetic analysis and conversation analysis have approached response cries, physical effort and affect in interaction. By looking at how *huh huh* is used in Finnish conversational contexts allows us to see both the sequential organization and phonetic variation in pitch, intonation, stress and other suprasegmental features. The subsequent sections present a sequential analysis of cases, where Speaker A produces a stand-alone *huh huh* turn followed by Speaker B repeating (more or less) the same lexical item in response to A's turn. The analysis shows that these micro-sequences work as an entry ticket or testing ground for collaboratively assessing an ongoing action and through phonetic variation it may be best described as an inference-making device by and for the members. Finally, some conclusions, implications and limitations of this analysis are discussed.

Response cries establish interactional moments for “feeling the same”

There is a large body of literature on both interaction studies and linguistics that takes its starting point from Goffman's influential work on response cries defined as “exclamatory interjections which are not fully-fledged words” (Goffman, 1978, p. 800). We utter such English cries as *phew*, *wow* or *oops* when we find ourselves “unable to shape the world the way we want to” (ibid., p. 818). For Goffman, response cries are devices for the speaker to momentarily step out of the conversation, while Schegloff (1997) refers to similar items as “pre-lexical grunts”, suggesting a move towards, rather than away from, lexical items. Further, recent literature suggests that at least some response cries are, in fact, proper lexical items (Dingemanse, 2017, p. 198) and this holds true also in the case of the Finnish *huh*. It belongs to a group of “consonant h + vowel + consonant h” structure items, each with a distinct meaning and context of use: *hih*, *heh*, *hah* and *häh* as varieties

of laughter, *häh* also as a repair-initiator, *höh* as disappointment and *hoh* (*hoijaa*) in connection with tiredness. Furthermore, the phonestheme *hu* can be found in Finnish words related to airflow, speed and surprise, for example, *hujahtaa* ‘to flash’, *hulmuta* ‘to wave’ (Jääskeläinen, 2013, p. 18) and possibly also *huoata* ‘to sigh’ (although the argument is here maybe valid semantically rather than etymologically and phonologically). Also grammatically, *huh* can be used as a recognizable word-like item, as Fragment 1 from a Finnish animated police sitcom “Pasila” shows. A female officer (A) is showing crime scene images to her fellow officers (including the police chief B) when her private holiday photos mistakenly appear on the whiteboard.

(1) Pasila

01 A ((*projects a personal photo on the white board*))
 02 → B: ↓No ↑huh ↓huh mikä friidu.
 PRT what chick
 NO HUH HUH what a chick
 03 ↑EIH E:I >FRIIDU<, =E:I >HUH HUH<,
 no not a chick not a HUH HUH

Line 1 marks the triggering event, an unexpected appearance of out-of-context photos. B inadvertently makes a public reaction (pleasant surprise) with the turn-initial particle *no*, a reduplicated *huh huh* response cry, and then closes the turn by describing the emotion-inflicting object as a “chick” (line 2). Scripted as realizing he has cried aloud an inappropriate comment publicly, B immediately self-repairs and denies the explicitly inferred categories of a “chick” and the *huh huh*, treating them as equally lexical items (line 3). Through self-repair, B’s “escaped control” (Goffman, 1978, p. 799) becomes even more observable. Scripted for a sitcom, the humor of this event is based on the recognizable machinery of *huh huh*: we come to see what B thinks the recipients might make out his response, a response that should never have been made public. We also note the suprasegmental stylization in *huh huh*: the response cry *no huh huh* (line 2) is

produced with a wide pitch span ($\downarrow no \uparrow huh \downarrow huh$) while in line 3 the pitch movement on the rapidly rushed through word-like *huh huh* is integrated with the rest of the utterance.

Most *huh huhs* in the data are not this type of flooding response cries, but are rather controlled recipient-designed turns in the sense that they are designed to prevent misunderstandings of the speaker's stance. They are intended for the benefit of others. Especially in transitions within a physical activity, speakers choose to produce *huh huh* rather than deep outbreaths (natural) or detailed accounts (saying), to borrow a conception from Wharton (2003) who discusses interjections along a continuum of communicative behaviors between 'showing' and 'saying'. While *huh huh* as a response cry makes speakers' inward state available to the recipients, it also serves as a device for achieving some activity, possibly other than merely showing the immediate response. Wiggins (2013), for example, argues that displays of disgust in family mealtimes are not merely responses to food: they make available moral judgements about the target object and eating practices within a social setting.

Participants engaged in the same activity typically have equal access to "the events, activities and sensations which [they are] entitled to evaluate by virtue of having experienced them" (Heritage, 2011, p. 160). In these cases, the shared experience should provide similar meaning of *huh huh* to all participants. As Goodwin (1996, p. 395) states, a response cry "establishes the unproblematic existence of an event and sets parameters for how it should be known". Yet, engaging in a shared activity offers only a slot for "feeling the same", that is, we have very few resources to know if both participants witness the event in the same way. We can question, for example, whether two runners have the same experiential access to the physiological state that makes a *huh huh* a legitimate cry.

What we have evidence for, however, is that an affiliating response to a response cry is highly preferred. Heritage (2011, p. 174) claims that the distinctive quality of response cries as vehicles for emphatic alignments *in the second position* is that they do not differentiate between the report of the event and the event itself and for that reason they “can attain a closer degree of empathy with the reported experience than might otherwise be the case”. By not discriminating between the recipient’s stance and the stance embedded in the storytelling, response cries evoke and claim a degree of empathic union and build a platform for launching more propositional and substantive forms of affiliation (ibid.). This study asks whether this is also the case with *sequentially paired response cries*. I will discuss cases where recipients offer a candidate understanding for the first speakers’ response cries in cases where participants have different degrees of access to the triggering event.

Lastly, there are cases in which the recipients seem deliberately to bypass one possible understanding of the initial response cry. Instead, they expand the sequence in a direction, where we, as analysts, find evidence of recipients disassociating themselves with the potentially problematic stance conveyed in the first assessment. By repeating the same response cry, recipients create a slot for re-framing the action. Couper-Kuhlen (2012) claims that when verbal devices such as response cries are used in producing affiliating response to a telling, there is an element of prosodic matching or upgrading while prosodic downgrading marks a disaffiliating response. This study suggests that maybe response cries also in the first position are phonetically designed to perform certain actions: to make the triggering event available to talk about, to call for an affiliating response rather than to treat it as something that needs a remedy. Additionally, if the recipient finds it problematic to respond to the response cry, then a simple affiliative move would be to repeat that response cry.

The data, the collection, and the methods

The data is from three video corpora of Finnish talk-in-interaction: 1) Human Activity in Natural Settings (HANS, University of Oulu), a corpus of mobile and social interaction in various outdoor settings (total 54 h); 2) the Poliisit corpus, a documentary television show following the daily working tasks of Finnish police officers (total 180 h); and 3) the Vain elämää (VE) corpus, a Finnish television music show (total 2 h).¹ The cases presented here from the HANS data were collected by the participants wearing a GoPro camera on a harness. For this reason, only the point-of-view of one participant is available, whereas the other participants can be seen on the recording while their voice may not always be available to the analyst due to their distance from the recording device. The data from Poliisit and VE was downloaded from a streaming service and consisted of the final television production. This provided an efficient way to add relevant cases to the collection from a particular contextual environment where emotional displays and assessments of observable events are frequent (Kidwell, 2009; Hutchby, 2006). Using response cries is one way of doing assessments, therefore allowing the study of members' methods of dealing with any ambiguity in their stance. These recordings also provide professional sound quality that facilitates auditory analysis. The obvious downside is that even documentary films and TV reality shows are subject to manipulation (Bartsch, 2014) in an effort to highlight and show the audience a particular interpretation of an actual encounter. The possibility of production stage 'frankenbiting' was considered particularly in the selection of cases analyzed in this article.

¹ In addition, Fragment 1 outside the three corpora highlights some aspects of *huh huh* that are not found elsewhere in the data. All research participants in the HANS corpus have consented to the materials being used for research, and their real names have been replaced in the transcripts. In the VE and Poliisit data, an explicit decision was made not to acquire consent from the participants because 1) the access to the recordings does not require registration to the streaming service, 2) the participants appear in the data as public figures, and 3) this study is about practices rather than individual participants (Legewie & Nassauer, 2018).

The collection for this study includes 34 instances of *huh huh* response cries. The criteria for selection include: 1) that there are at least two participants present: a speaker and a recipient (hence instances of self-talk are excluded); 2) that the opening turn in the sequence comprises of a *huh*-related lexical item alone or as part of a turn (cases where *huh* is used in response to a telling are excluded); and 3) that an action or turn preceding and following the target turn are available for analysis (this criterion excludes some occurrences of the phenomenon in TV production where there is a cut in the recording prior to the relevant turn). Based on a descriptive categorization of the collection there is some variation in terms of the lexical form of interjection, its position within the turn and its sequence. The majority of the cases in the data are reduplicated stand-alone *huh huh* turns in the first position.

This study deals with the vagueness of language from a conversation analysis point of view. Rather than trying to intuitively describe or categorize what type of affectual state of mind the speaker might have had when uttering interjections like *huh huh*, conversation analysis focuses on the treatment of those utterances as actions taken by the co-participants in terms of their observable actions (see Curl et al., 2006, p. 1725). The sequential organization of talk-in-interaction lays the foundation for the analysis: rather than mapping out formational and suprasegmental variations in the realization of *huh huh* per se, the analysis is attentive to the use of the item for interactional purposes. The role of phonetic analysis focusing on F0 contours is to help understand any possible similarities and differences in the ways in which *huh huh* is used to mark access to or the degree of affectual display.

Phonetic analysis

Previous conversation analysis studies in conjunction with phonetic analysis testify to a close relationship between the action conveyed in a turn and its phonetic format (Wilson & Wharton, 2006). This concerns incitement (Reynolds, 2016), second assessment turns (Ogden, 2006), responses to proposals (Stevanovic, 2012), third turn responses (Sikveland, Skovholt, Stokoe, & Skarbø Solem, 2018), and formulaic patterns such as summons (Sikveland, 2019) and repair-initiations (Curl, 2005; Curl et al., 2006; Walker & Benjamin, 2017). Response cries, for example, may be separated from the rest of the talk through pitch, stress and other paralinguistic effects, which seems to be a natural observation given the strong affectual ingredient. Yet, labelling *huh* *huh* either as a display of surprise or fatigue only based on its phonetic qualities is highly subjective. In addition, same response cries can be used to show both affiliation and disaffiliation to, for example, complaints and stories (Couper-Kuhlen, 2012). Therefore, as Ogden (2006, p. 1753) states, any strong claims about universal or categorical labels should be warranted; evidence of a phonetic format should be viewed from the participants' displayed orientation to the talk and actions performed in a context through the talk.

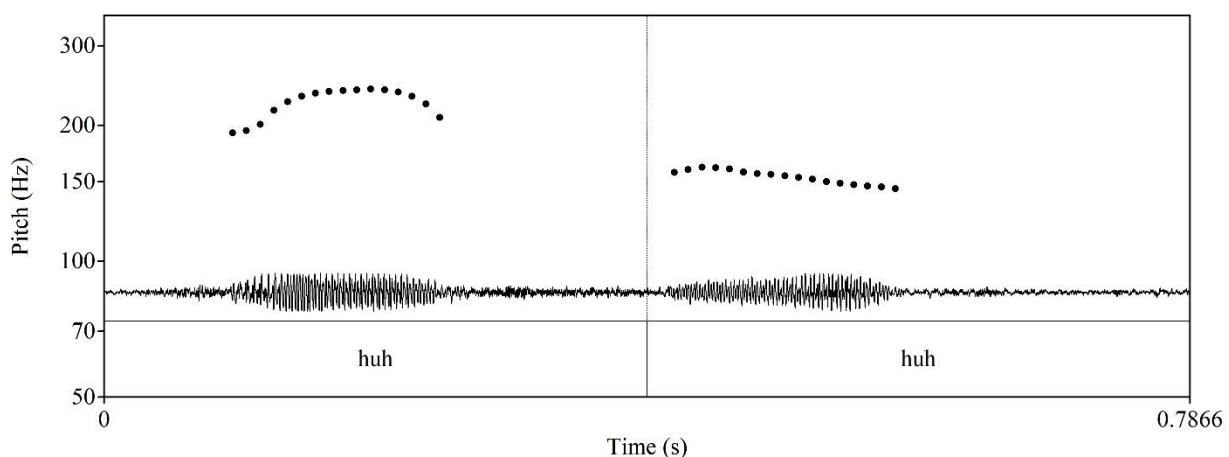


Figure 1. The canonical phonetic format of Finnish *huh huh* (from Fragment 2).

Having said that, there is a canonical pitch pattern in reduplicated *huh huhs*, shown in Figure 1 presenting both the F0 trace (plotted on a logarithmic scale) and the waveform generated using

PRAAT. In cases where the audio data collected in the wild was subject to “noise” from the environment and did not provide an acoustic record, an auditory analysis was used. *Huh huh* [^huh^w ^huh^w] is realized with close back vowel [u]: starting with gradual breathy onset (a voiced glottal fricative [^h]) through a final fricative [h], where the dominant friction is typically labial. There are cases where the final friction is glottal, velar or pharyngeal. In all cases except one, there is an intonational downstep from the first to the second syllable [̄huh ↓̄huh]. This format can be varied particularly in terms of rate of articulation, pitch, the semitone difference between the syllables and the breathiness. The phonetic properties for the data fragments are presented in Table 1.

Table 1. Duration and pitch span (in semitones, measured from the midpoints of the two syllables) of Fragments 2–4 and 8–9.

	Fragment 2	Fragment 3	Fragment 4	Fragment 8	Fragment 9
	↑ <i>huh hu(hh)h</i> .	↑ <i>huh</i> ↓° <i>huh</i> °.	<i>hu(h)h</i> ↓ <i>hu(h)h</i> .	> <i>huh</i> ↓ <i>hu(hh)h</i> .<	> <i>huh</i> ↓ <i>huh</i> .<
Duration (s)	0.70	0.66	1.17	0.60	0.47
Semitone difference	-7.6	-10.6	-5.2	-7.2	-4.2

The monosyllabic lexical unit *huh* is typically reduplicated (*huh huh*) but combinations with other lexical items are also found in the data: *no + huh huh* and *huh + huijaa/heijaa*. The collection includes no cases where particle *no + huh (huh)* would be used as a stand-alone in first position turns; this is because *no*-particle typically connects the turn to the previous turn-in-talk (Sorjonen & Vepsäläinen, 2016) although it can tie turns further backward across considerable time spans (in

Estonian, see Keevallik, 2013).² Presumably adding items before *huh huh* distances the item from its natural relation to breathing. Therefore, if speakers are to express their state of “being out of breath”, the way would be to start with an exhaled *huh*. *Huh* can be combined with items like *huijaa* or *heijaa* to produce idiom-like expressions that are one step further toward standard language than a *huh* would be.

In some cases, there is no audible gap between the two parts (thus, *huhhuh* would be an appropriate orthography), while in most of the cases there is a clear exhaled part between the syllables (thus, *huh huh*). In reduplicated *huh huhs*, the latter *huh* can be produced extensively by exhaling, as is the case in Figure 1 ([^huh:^w ^huh^w]). This reinforces the notion of *huh huh* as a conventionalized version of outbreath or sighing (Hoey, 2014) rather than a response particle (see Pehkonen, forthcoming). We can thus depict *huh huh* as a conventionalized version of a natural expression.

Sequential analysis: making a response cry publicly noticeable

Emotion in interaction is a sequential phenomenon and turn-taking can be used to share emotions (Stevanovic & Peräkylä, 2015). A look into the data shows that *huh huh* also works as an affective display that builds on a preference for both an aligning and affiliating response and therefore helps organize the social interaction (see also Pomerantz, 1984; Stivers, 2008). When a speaker publicly produces a stand-alone *huh huh* in the first position with a falling intonation contour, and leaves it there, the recipient should produce the next turn. This next turn is the key to analyzing what type of action the recipient thought the speaker was doing with their response cries. As the analysis shows,

² In Fragment 1, *no huh huh* was a response to the triggering event but as a part of a syntactical TCU construction.

the recipient typically produces an aligning and affiliating response. In Fragment 2, two police officers are getting into the car after having dealt with an aggressive customer.

(2) pol_huh_28

01→PO2: ↑hu(h)h hu(hh)h.

02 PO1: Tässähän pääsee kohta niinku (.) mukaan (0.3) meininkiin.
 here+CLI get soon PRT along action
 One is about to get into the swing of things here

The termination of previous strenuous physical activity triggers a *huh huh* from PO2 (line 1)³: the first syllable produced with a pitch higher than his normal level and with a clearly audible outbreath. Although the stand-alone *huh huh* gives few cues about the type of action it proposes, PO1 has no problem in producing a contextually relevant next turn, which is an assessment-like turn that continues the action initiated with PO2's response cry. The grammatical evidence for an affiliative response is in the clitic particle *-hän* that shows PO1's access to the same information (VISK §830) which is necessary to recover the intended interpretation of PO2's preceding turn.

Huh huh initiates a shared emotional display also in Fragment 3. Here, a police officer P2 is waking up an intoxicated citizen CIT. Up until line 9, P2 has tried to wake CIT up by giving him multiple directives without a sufficient response from CIT.

(3) pol_huh_26

09→CIT: ↑huh ↓°huh°.

10 PO2: £Ä(h)lä muuta sanof.
 don't anything else say
 My thoughts exactly

³ The example also shows that speakers can manipulate the transition time between the triggering event and talk-in-interaction about that event. That is, while the reaction is assumedly imminent, the response can be delayed until a moment when talk becomes possible.

CIT's *huh huh* in line 9 provides a kind of vocalized self-assessment that replaces what could be expected as a more conventional opening turn for someone waking up, such as "good morning".

Huh huh seems to provide a particular kind of entry into a focused encounter—one that recognizes the atypical nature of the encounter. Again, the recipient P2 has no problem in aligning with CIT's response cry. PO2's response (line 10) formulates the shared experience of CIT coming back into this world: *my thoughts exactly* suggests that CIT has successfully expressed what anyone could have said in this situation. The response cry thus performs the dual action of entering into discussion through a stance that the recipient can align with.

In order to produce immediate affiliating response to *huh huh*, both speakers need to have the same experiential access to the triggering event. In both previous cases, the first speakers' *huh huhs* are displaying a stance toward a reciprocally perceptible event. This is in contrast to previous studies where the reaction source is to be found in the preceding turn-in-talk (e.g. Wilkinson & Kitzinger, 2006) or where the second position turn affiliates with a storytelling (Stivers 2008; Couper-Kuhlen 2012). We might suspect that cases like above can only occur when there is a focused interaction going on. In unfocused interaction, *huh huhs* like these might go unnoticed, unless extra work is done to make a response cry publicly hearable. Speakers can do this by providing sufficient time and new transition relevance places for the hearer to respond to a response cry. The sequence in Fragment 4 is triggered by the police patrol's ongoing (unsuccessful) task of searching for a troublemaker in a very quiet neighborhood. PO1 is driving, while PO2 is glancing around.

(4) pol_huh_10

01 →PO1: @hu(h)h ↓hu(h)h.@ ((yawns))

02 (2.4)

03 On tämä rikollista sakkia (.) ja rikollista seutua.
is this criminal crew and criminal neighborhood
This is a really criminal crew and criminal neighborhood

04 PO2: On.

it is

PO1 produces *huh huh* through a yawn (line 1), followed by a 2.4-second pause (line 2). PO2 does not produce a response. While we have no access to what PO2 makes of PO1's *huh huh* (he may have simply heard it as a verbalized yawn), we can see that with the verbal extension concerning the environment (line 3), *huh huh* explicates PO1's stance toward their task. PO1's reference to criminality becomes heard as an expression of irony, something that PO2 can readily recognize with his affiliating response *it is* (line 4).

By being underspecified in their nature, stand-alone response cries risk being insufficient for the recipient to recognize. For this reason, explicit verbal accounts can be provided within the same turn, as is the case in Fragment 5. During their lunch break, PO2 enters an outdoor patio located in the backyard of a police station, while PO1 is already seated and eating.

(5) pol_huh_15

01 PO2: → no ↑↑HUH, (.) täällähän on ihan ku etelässä.
PRT here+CLI is quite like south
NO HUH this feels just like in the south
02 PO1: nhm.

Here, the response cry *no huh* displays both PO2's perceptual response to stepping outside on a hot afternoon and her entry into a casual discussion about the weather. PO2, being already outside and thus exposed to same external condition, has, in principle, the same access to the referent of PO1's assessment. However, because of the recipient's *being already* there and thus not himself undergoing the transition, a further explanation of PO2's response cry is needed. In fact, we may notice that PO1's response, while being rapid, is minimal and shows a low level of engagement. The response cry has nevertheless launched a sequence, even though a short one.

The placement of *huh huh* is indexical to the ongoing action but not necessarily in the same way that a response cries such as lifting grunts (Reynolds, 2017) are—that is, the production of *huh huh* as a reciprocity-seeking item can be placed so that it can be heard in relation to a relevant triggering event which makes a subsequent assessment relevant in the next slot. For example, in Fragment 2, the officers wait until the source of trouble event (their customer) cannot hear their conversation. Finally, it should be noted that not all response cries result in an extended assessment sequence. Keevallik (2018) has proposed that efforts to invite responses from other participants to the first speaker’s observations of the surrounding environment differ from self-talk in that incipient talk is carefully timed to the available participation framework. This seems also the case here: the *huh huhs* that receive a response are located so that the recipient can produce a response.

***Huh huh* as an inference-rich transition ticket to focused encounter**

An exemplary setting where emotional displays followed by verbal assessments of a triggering event can be found in the Finnish version of the international reality television format *Vain elämää* (The Best Singers, lit. *It’s only life*) where a group of musicians come together over a dinner and perform each other’s songs. The format builds on positive assessments and affiliation: Artist A turns something initially owned by Artist B into something A-like and then B is expected to react to A’s “remake” emotionally. In this type of setting, the song as a triggering event is available to all the participants, but some know it better than the others do. A collective response (with possibly some clearly marked individual response cries) is not enough but an exchange of gratitude between the Artists A and B is needed. The program therefore uses a script: when a performance ends, Artist A leaves the stage while Artist B stands up to welcome Artist A. At this point, these two participants establish an interactional moment for the mutual exchange of emotional displays. This

conversational slot is ideal for a variety of emotional displays such as laughing, crying, being observably “speechless”, or, in fact, for response cries like *huh huh*.

The next two data fragments feature Arttu as the recipient artist whose songs Aki (Fragment 6) and Danny (Fragment 7) are performing. The reader should be attentive to how performers (Aki and Danny respectively) and Arttu establish the interactional moment for expressing emotions and gratitude. The relevant lines are marked in the transcripts.

(6) Vain elämää_3

01 ((AKI's performance ends))
02 ART: va:::u.
03 ALL: (4.0) ((clapping))
04 ART: → ↑↑huh.
05 (2.3) ((clapping))
06 AKI: >he he he<
07 ART: hhh
08 SAN: okei.
Okay
09 +(0.7) ((clapping))
+((ART & AKI stand up & walk twd each other))
10 AKI: → >huh huh<.
11 (0.5) ((clapping))
12 ART: → ↑huh ↓heijaa.
13 (2.0)
14 ART: heh %[heheh he
%hugs AKI
15 AKI: +[melkonen biisi.
quite song
one helluva song
+hugs ART
16 (0.6)
17 ART: kiitos.
thanks
18 (0.7)
19 fihan loistava vetof.
really excellent show
20 AKI: .pHHHhhh

15 (0.3)
 16 SAN: **sill[ä lailla] sitte.**
that's the way then
 17 ART: **[kiito-]**
thank-
 18 (1.3)+%
 --->+
 --->%
 19 DAN: **olet tehnyt hienon teok[sen.**
you have made a great piece of art
 20 ART: **[.khiitos.**
thhanks

There are similarities but also differences between the two cases. While the target lines (line 10 in Fragment 6 and line 10 in Fragment 7) are not the first response cries nor turns produced, they are the first turns produced as an entry into a focused encounter. Thus, they provide an entry for the ratified participants to proceed from choral audience responses towards more intimate ones. In both cases, the pair is followed by expressions of gratitude and explicit assessment about the quality of the song performed. The two main differences here concern who initiates the sequence and whether the second repeated response upgrades or downgrades the initial response cry. The former is perhaps to be expected, given that the situation is designed to be emotional for both parties and both parties are institutionally required to express their gratitude.

The two repeated *huh huh* pairs might just nicely capture a more general problem of expressing and assessing gratitude (Robles, 2012). In Fragment 6, Aki's rapid *huh huh* (line 10) works as a sigh of relief as if "having barely made it", while Arttu's response *huh heijaa* (line 12) does two things: by repeating the *huh*-initiated TCU, it aligns with the breathtaking-ness of the performance, but by lexically and prosodically upgrading the turn, it disaligns with Aki's potentially dismissive self-assessment. The prosodic upgrade is realized through the stressed first syllable of *hei-jaa* and the large downstep: the 14.0 semitones difference between the *huh* and *heijaa* is the largest one found in the data.

In Fragment 7, the high pitch and downstep in Arttu's *huh huh* (line 10) marks the turn as conveying a strong emotional stance, compared with his first 'flat prosody' *huh huh* (line 7) produced within the joint audience response sequence (lines 2–9). Line 10 provides therefore an overtly positive appreciation of Danny's performance, while Danny's *huh huh* (line 12), with its exceptional rising final intonation, could be heard as doing something close to repair-initiation. Both examples show ambivalence towards who or what should be assessed: the performance or the song (see Danny's assessment in line 19). In sum, *huh huh* works as an entry and exit ticket between sequences and actions.

***Huh huh* is followed by a recipient providing a candidate understanding**

In the previous two fragments, emotional displays are observably expressed in a setting where everyone knows that *huh huh* is a response to the musical performance and where some participants are more ratified to express their emotions due to the roles given in the script. The roles are not always that clear, however. Particularly when participants are engaged in shared physical activity or when the triggering event is one among several possible events, both the speakers and recipients may design their turns to maximize the possibility for intersubjective understanding. On the one hand, we already saw in Fragments 4 and 5 that when the speakers initiated their turns with a response cry they extended or finished their turns with a verbal elaboration, thus saving their recipients from merely guessing what type of action the speakers might have intended. On the other hand, the recipients can provide candidate understandings for speakers' response cries. These are often formulated as polar questions, with a preference toward a yes-answer. This is the case in Fragment 8 where dad has been picking berries in the forest with his three small children. The

Fragment starts with Anni responding to dad's suggestion that they go home: she has played the game called "the gate", where anyone passing her has to go under her extended arms (line 1).

(8) HANS30_GP020785

01 ANN: ja sitten porttikin lähti +kotii::n?=
and then gate+CLI went home
and then also the gate went home
dad +climbs on a cliff--->

02 DAD: =sitten kiivetään ylöspä:in?
then climb up
then we climb up

03 (1.4)

04 ENN: uh uh.

05 ELL: un tä:..+
dad --->+stops, head turns gazing Enni--->>

06 (4.6)

07 ENN: uh.

08 (2.5)

09 → DAD: >huh ↓hu(hh)h.<

10 (1.3)

11 ENN: e:h (1.0) ftuliks sul↑la:£ (1.1)
became you
erm did you get

12 kuuma ku sä kiipesit ton vua°ren°.
hot when you climbed that mountain
hot when you climbed that mountain

13 (1.0)

14 DAD: no (.) muun muassa siitä.
PRT among other things that
well partly because of that

They climb a cliff which turns out to be a strenuous effort, particularly for 6-year-old Enni moving on her own (note the grunts she makes in lines 4 and 7) and dad who needs to physically assist Anni (4) and Ella (2) to get up the cliff. After lifting the girls up, dad produces a response cry *huh huh* with an exhale in the end (line 9). After a pause (line 10), Enni initiates a turn with hesitation marker *e:h* (line 11). She then provides her candidate understanding of dad's *huh huh* by combining contextual information on what could reasonably be the triggering action (*climbing a mountain*) with a possible physiological state (*getting hot*) that could warrant a response cry (lines 11–12). Dad

provides a partial confirmation with the use of turn-initial *no*-particle marking a departure from the expectations set in the previous turn (see, Sorjonen & Vepsäläinen, 2016, p. 259). His answer affirms the *huh huh* is to do with his physical state but not necessarily confirming that climbing is the only reason for his fatigued state.

Norrick (2007, p. 167) states that “the pragmatic functions interjections realize seem always to be clear to listeners in the concrete context”. In the data there are, however, cases, where the recipient’s candidate understanding seems to fall short, resulting in the recipient missing the action initiated with the response cry. In Fragment 9, two police officers are engaged in a task of escorting an aggressive customer to a hospital. PO2 travels in an ambulance, while PO1 follows the ambulance in their police car. The fragment starts when PO2 steps down from the ambulance.

(9) pol_huh_4

01 PO2: +↑opsaa::.=
+steps down, smiling

02 PO1: =Matka suju mut[kitta,
trip go smoothly
The trip went smoothly

03 → PO2: [>huh ↓huh< £↑he↑he↑he£

04 PO1: £Oliks siel lämmin vai£.
was there hot or
Was it hot inside there or

05 PO2: £No e::i£ [(-)
PRT no (-)
Well no (-)

06 PO1: [Eiks teil ilmastointii
no-Q you air conditioning
Don't you have air conditioning

07 oo ambulans[sis.
have ambulance
in the ambula[nce

08 PO2: [@o::: [::n@.
Yes we do

09 EME: [o::n.
Yes we do

10 (0.5)

11 PO2: Paljon pa(r)e:mpaa kyytii ku meitä £paril.=
much better ride than our couple
The ride was much better than in our car

12 =No e::i vaanf.
 PRT no just
 =No just kidding
13 Ha ha ha ha A::H

PO2 starts the encounter with a playful grunt (line 1) while stepping down from the ambulance. PO1 immediately launches a task-related inquiry (line 2) which is, however, interrupted by PO2 who produces yet another response cry, a rushed through *huh huh* followed by a burst of giggles (line 3). PO1 abandons his first effort to gain information from PO2 and instead produces a candidate understanding of the circumstances that might have triggered PO2's *huh huh* (line 4). PO1's suggestion is identical to the one provided in Fragment 8 (line 11) and is certainly valid, given the prevailing warm weather. It is also produced in a laughing voice and in a question form but with no hesitation (in fact, in line 6, PO1 still pushes his initial understanding as situationally valid). Furthermore, PO2's response to PO1's candidate understanding starts again with the *no*-particle, which here not only marks the suggested understanding as out of the question but also alerts the recipient that something else than affirmative response is on the way (Sorjonen & Vepsäläinen, 2016, p. 260). From line 10 onwards, PO2's initial trajectory—*huh huh* as a precursor to jocular chitchat—is made explicit.

In sum, recipients can provide their candidate understandings of the first speakers' response cries, revealing participant analysis of what such a cry might mean. One reason for this is that recipients do not always have direct access to the triggering event. If speakers resist being seen as initiating a sequence or action of certain type then the social solidarity between the participants may be temporally compromised (as was shown in Fragment 9 and will be discussed in the following section).

***Huh huh* is shared but action is reframed**

It would be only a partial inquiry to look at response cries as backward-looking reactions to a triggering event. Because of the transitory role of response cries, *huh huhs* reach back to something immediately experienced in order to perform a next action. Fragment 10 shows that while determining the reason for and the meaning of the initial *huh huh* is unproblematic to the recipient, the response turn (*huh heijaa* in this case) might not only express affiliation but also something extra. The recipient can utilize the inference-rich nature of *huh huh* and reframe the action performed in the extended sequence without making the disalignment too obvious.

Two friends, Pete (wearing a GoPro) and Ollie, are running an orienteering course together. What we can tell from their overall comments during the event is that neither of them enjoy running. Pete is setting the pace. At the point Fragment 10 is taken from, Ollie has been starting to lag behind Pete. Of analytical interest here is how Ollie's initial displays of his struggle are reframed into social solidarity and a preference for keeping up with the running.

(10) HANS16_GP010700

```
01      +% (3.0)
      pet  +>>decelerates and starts walking--->
      oll  %>>runs behind PET--->
02  PET:  +huh.          +
      +glances at OLL+
03  OLL:  %aah:: .hh hhhhhhhh .hhhhhh hhhhhh
      %walks--->
04  PET:  Herralletta puol tuntii juostu (randoria) mutta tämmöne näin.
      INT          half hour   run      (randori) but   PRT    PRT
      Goddammit been running (randori) for half an hour but like this
05  OLL:  Jaa (HH)h .hhhhhhh hhhhhh .hhhhhh
      I see
06      .hhhhhhhhh hhhhhhhh .hhhhhhh hhh .hhhhhhhhhhh
07 →    <↑hu (hh) h ↓hu (hhhhhhh) h.>
08      (0.8)
09 → PET:  ↑hu (h) h ↓heijaa.
10  OLL:  a::rgh.
11      (1.0)
```


12 PET: ↑**Hauskaa tämä on.**
 fun this is
 This is fun

13 **(0.5)**

14 OLL: **Kyllä tää niinku normilenkkeilyn voit(h) taa.**
 sure this like normal jogging win
 This surely beats normal jogging

15 PET: **@Jo vain helevetin [hyvin.@**
 sure hell well
 Surely hell yeah

16 OLL: **[h(h)eh hee .hh hhhhh .hhhhhhh hhhhhhh**

17 **.hhh+hhhhhh %HHHHHHHHHH .(n)hhhhhh ä: :hhhhhhhhhhhhhhhh**
 pet --->+starts running>>
 oll --->%starts running>>

Pete starts to walk when, from a navigational point of view there is no need to go slowly (line1).

This may be because of his fatigue, but for Ollie Pete’s change of pace offers a chance to catch up with him. Pete quickly glances over his shoulder and produces a stand-alone *huh* (line 2). Ollie reaches Pete, starts walking too and produces a relief grunt of *aah* (line 3). Pete then produces an assessment (presumably) comparing running with free-style practice in martial arts (randori) they are both familiar with (line 4). Ollie responds with a neutral response token *jaa* (‘I see’, lit. ‘yeah’) (line 5). After extensive inhaling and exhaling Ollie produces a clearly marked *huh huh*, a response cry which audibly exhibits his physical discomfort (line 7). After a short pause (0.8), Pete responds to Ollie’s turn by repeating a slightly modified response cry *huh heijaa* (line 9).

By repeating the previously issued lexical items, speakers decline to further the trajectory of the preceding sequence (Curl et al., 2006, p. 1733). If we look at the phonetic elements in Ollie’s and Pete’s response cries, however, we could claim that Pete in fact does something extra with his repeat. Ollie’s *huh huh* in line 7 is phonetically typical for an exhaled stand-alone reduplicated *huh huh* where the first syllable has a higher pitch than the second. However, the semitone difference between the two syllables is relatively large, 9.0 semitones (from 282 Hz of the first *huh* to 168 Hz of the second measured from the midpoints of the two syllables). The response cry lasts a long time (1.15 s) and there is a clear exhale between the two parts (cf. Hepburn, 2004 on breathiness as a

component of emotional displays in crying). These properties and the contextual environment mark Ollie's response cry with strong display of fatigue. If we compare this with Pete's responsive *huh heijaa* turn, we see both items sharing a somewhat similar F0 contour (Figures 2 and 3) in the first part (*huh*) but the latter part is different in terms of both the lexical item (*huh* → *heijaa*) and the phonetic features: the semitone difference in Pete's turn is 5.3 compared to 9.0 in Ollie's turn.

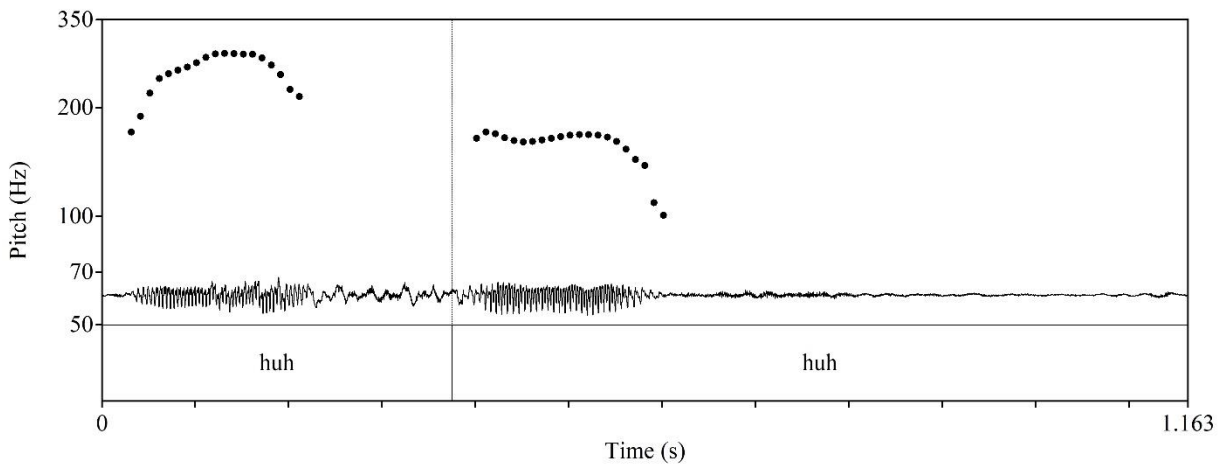


Figure 2. Ollie's *huh huh* from Fragment 11, line 7.

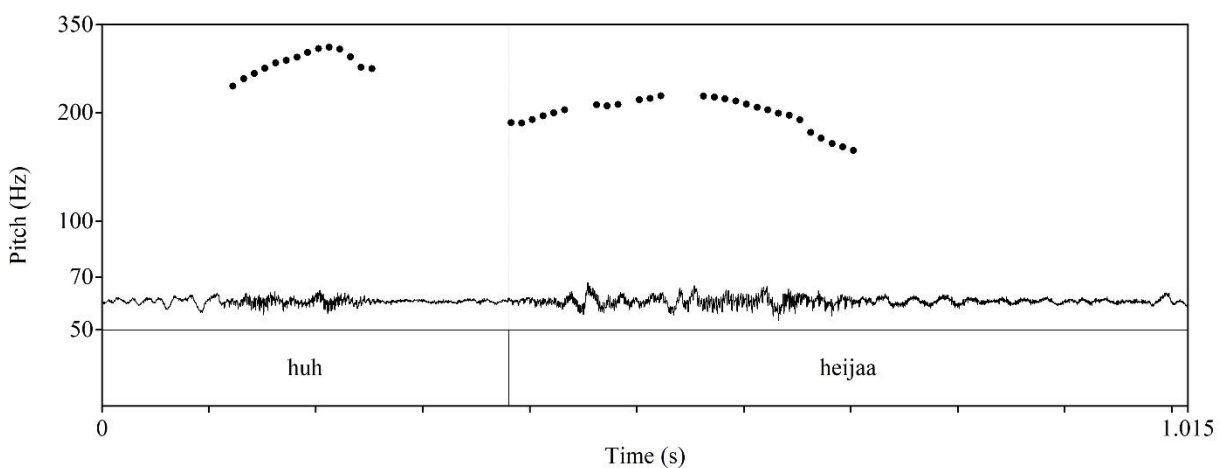


Figure 3. Pete's *huh heijaa* from Fragment 11, line 9.

Fragment 6 included a similar sequentially paired *huh huh–huh heijaa* segment, where the phonetic upgrade (through stress and downstep) was used by the recipient to show disalignment with the stance of the first speaker. Here, the *heijaa*-part of Pete’s response (line 9) is produced with no stress and with a rather modest downstep and it is hearable as aligning rather than disaligning in relation to Ollie’s previous turn. Lexically the turn is modified by replacing the reduplicated *huh huh* with a three-syllable construction *huh + hei-jaa*. The *heijaa*-item derives from the verb *heijata* (to swing), which makes the modified ‘airy’ second assessment depart from the ‘exhausted’ physiological response cry. Pete does not expand his turn right away. By not producing any further talk, he also declines to proceed with the sequential structure. What happens next is that Ollie produces a vocalization *argh*, an open vocal tract pharyngeal or epiglottal fricative (line 10). Whether this vocalization is primarily somatic or social in nature remains unanswered but what it does in terms of turn allocation is that it fills the slot for the next turn (third position assessment; Schegloff, 2007, p. 118). We can thus see a sequence where Ollie’s first position *huh huh* marks his discomfort with running and places Pete in a position to respond; Pete’s second position *huh heijaa* modifies and only partly aligns with the possible action trajectory of expressing discomfort; and Ollie’s third position turn repeats the physical discomfort and stance delivered in his first turn. This puts again Pete in the position of the next speaker.

After a 1.0-second pause, Pete produces a verbal assessment *fun this is* (line 12). With that turn, Pete’s non-alignment with Ollie’s initial response cry becomes noticeable. Working the sequence backwards (as analysts), we can see Pete’s *huh heijaa* has a more elaborate explanation. Instead of emphasizing his discomfort with running, Pete’s turn in line 12 transfers the sequence toward a positive assessment. The ordering of words in the sentence, starting with the word “fun”, make the turn appear to be a repair to any possible misunderstanding that his *huh heijaa* might have caused and certainly to Ollie’s previous turns. More importantly, Pete offers a positive assessment of their

collective action for Ollie to align and affiliate with. Indeed, Ollie produces as strongly affiliating response (*this surely beats normal jogging*), starting with the work *kyllä* ('sure') which is epistemically reinforcing in a turn-initial position (see, Hakulinen, 2001; Keevallik & Hakulinen, 2018) and produced with laughter (line 14). However, Ollie still maintains his initial stance by contrasting their ongoing running action with something (apparently) less fun, namely "normal jogging". With this reframing, Ollie's turn receives a strongly affiliating response from Pete (line 15). The sequence ends with overlapping laughter highlighting the collaborative nature of their assessment. Given the start of the sequence marking running as problematic and closing the sequence with affiliative laughter—and indeed Pete starting to run without any further warning—shows that what is at stake is not the truthfulness of their assessments, but an effort to maintain the social solidarity through laughter.

Fragment 11 occurs somewhat later in the same orienteering course. Pete and Ollie are walking after having negotiated their next move on the course: they are to divert to a narrow path. Once they locate the path, Pete starts to run and the following sequence unfolds.

(11) HANS16_GP010700

```

01  OLL:  hhhh .hhhhhhh hhhh .hhhhh [hhhhhhh=
02  PET:                                     [+@ (hu: :) @
                                           +runs>>
03  →OLL:  =.hhhhhhh ↑hu(h)h ↓huh %herra siunaa(hhh) .
                                           lord bless
                                           HUH HUH Jesus Christ
                                           %runs>>
04                                     [. (n)hh hhhhhh .hhh]
05  →PET:  [↑↑HUH HU(hhh)H.      ]
06  OLL:  hhhh .hhh hhhhhh .hhh (h) argh .(n)hh

```

As Pete starts to run, he produces a one-syllable response cry, possibly a *hu(h)* (line 2). Ollie takes a deep inbreath and produces a quick (0.50 s) response cry *huh huh* combined with another response cry of *Jesus Christ*, before starting to run (line 3). At this point, Pete takes a quick glance over his shoulder at Ollie and repeats Ollie's *huh huh* but now with clear upgrading elements (line 5). Pete's *huh huh* is notably higher in pitch and it has distinctive phonetic qualities. It is produced at a pitch rate of over 400 Hz compared to his average F0 of 250 Hz and it has a 'pressed' voice quality, possibly accompanied with pharyngeal friction. This phonetic upgrade may convey the shared physiological experience but with a jocular twist. Pete's *huh huh* (line 5) treats Ollie's struggle with running as something Ollie is entitled to express but that should not compromise the progressivity of their shared running activity. In fact, neither Pete nor Ollie produces further turns, the lack of which is not treated as problematic.

In Fragments 10 and 11, the transition from running to walking and vice versa provides an interactional moment to express fatigue and to negotiate the consequences of "how that feels" for "what we should do next". When the participants are engaged in physical activity, both A and B 'own' their own physical bodies. They run individually, but their running takes place in a reciprocal interactional setting where not only "how A feels about doing X" and "how B feels about doing X" is important but rather: "how A/B feels about doing X with B/A as a way of doing X". Because of A cannot exactly know B's physical condition and vice versa, it seems reasonable to claim that the response cries are an important means to gain access to another party's stance or even their bodily experience and to establish preference for the continuation of the joint activity. As Goodwin (1996, p. 393) suggests, response cries are "organized as social phenomena that provide very powerful resources for shaping the perception and actions of others". If the participants are to fulfill a task together and maintain social solidarity, then the recipients are highly obliged to respond to a *huh*

huh. A relatively safe way is to repeat the response cry and to gain further information on how the other party feels and, perhaps, how we should feel together.

Table 2. Duration and semitone difference in paired *huh huh*s (Fragments 6–7 and 10–11)

	Fragment 6 line 10 (first) > <i>huh</i> <i>huh</i> .<	Fragment 6 line 12 (second) ↑ <i>huh</i> ↓ <i>heijaa</i> .	Fragment 7 line 10 (first) ↑ <i>hu:h</i> ↓ <i>hu(h)h</i> ,	Fragment 7 line 12 (second) <i>huh</i> ↑ <i>huh</i> ?	Fragment 10 line 7 (first) <↑ <i>huh</i> ↓ <i>hu(h)h</i> . >	Fragment 10 line 9 (second) <i>huh</i> ↓ <i>heijaa</i> .	Fragment 11 line 3 (first) ↑ <i>huh</i> ↓ <i>huh</i> .	Fragment 11 line 5 (second) ↑↑ <i>HUH</i> <i>HUH</i> .
Duration (s)	0.38	0.93	0.60	0.93	1.15	1.02	0.5	0.78
Semitone difference	-7.1	-14.0	-5.5	-5.7	-9.0	-5.3	-12.65	n/a

While it is impossible to claim that some given phonetic quality alone would explain the meaning of any single *huh huh* response cry, it is clear that in all four paired *huh huh* turns the repeated part did some upgrading or downgrading work either through lexical variation or phonetic quality (Table 2). In Fragment 6, the first *huh huh* was produced quickly, as if vocalizing a sigh of relief, while the response *huh heijaa* upgraded the first cry, and hence reframed the stance towards something positive, showing gratitude. The three-syllable item *huh hei-jaa* allows a greater semitone difference in pitch between the items. The same lexical modification was found in Fragment 10, but the reframing of action was the opposite: the *huh heijaa* (line 9) was prosodically matched with the *huh huh* (line 7) in terms of its length, and the semitone difference *huh heijaa* was even smaller than in the *huh huh*. All this suggests that the *huh heijaa* was used to show alignment with the initial response cry.

The data included two cases where the initial *huh huh* was repeated in an identical lexical form (Fragments 7 and 11). What we witnessed there, however, was a clear difference in the prosody and the phonetic qualities between the first and second responsive turns. In Fragment 7, the recipient produced a *huh huh* with a rising intonation and the first syllable lower in pitch than the second syllable, while in Fragment 11, the second syllable was produced markedly high pitched. Both these sequences were closed after the second assessment and for us as analysts there were no further turns to examine in search for evidence of what type of action they were doing. The participants, however, treated these minimal sequences as unproblematic for their ongoing activity.

Conclusion

This article started out with a motivation to look at how response cries bring together ongoing or closing social actions and assessing those actions. The claim was made that by studying what seems like a minimal and marginal sequence—in its fullest form Speaker A producing a response cry repeated then by Speaker B—allows us a fully-fledged analysis of how the participants create social solidarity. The study demonstrated that *huh huh*-initiated assessments are recipient-designed in terms of both their timing (vis-à-vis an ongoing activity), shared intersubjective stance (making the connection between an action and the depiction of that action observably hearable for the recipients) and what type of response is preferred. Occasionally, a stand-alone *huh huh* is enough for the recipient to recognize the action; sometimes a verbal elaboration is needed. It may also be that the recipients fail to acknowledge the triggering event altogether. In most cases, the recipients had no problem in producing a relevant response. When the response was a repetition of the response cry from the previous turn, the participants found this response to be appropriate and unproblematic (apart from Fragment 7's line 12 that could be heard as initiating a repair). Repeating the previous turn with "levelling prosodic features" is a safe way to maintain social solidarity. This finding

confirms previous studies (Curl et al., 2006; Couper-Kuhlen, 2012) on the relevance of phonetics and repetition: with prosodic downgrading or upgrading, the recipient can achieve something extra, such as displaying affiliation rather than treating the triggering event as something in need of a remedy.

Given the limited number of repeated cases of *huh huh*, it is fair to say that the phenomenon in question is real, but more research—including other settings (e.g., when feeling full after a meal) and vocalizations, possibly also in other languages for comparison—is needed to establish whether any phonetic generalizations can be made. This study also bears some obvious technical limitations. Having only one participant wearing a recording device with in-built microphone offers a partial representation of the audiovisual field available to the participants in situ. This may skew our understanding of what eventually should be analyzed as the first response to the triggering event. Studying social interaction in mobile settings would benefit from more technically sophisticated recording devices (ambisonic microphones and 360° cameras) than the one available in the current study.

Transcription conventions

<u>underlining</u>	– emphasis
HUH	– strong emphasis
-	– truncation
[]	– overlaps; timing of embodied action
=	– latching of turns
(0.5)	– pause length in tenths of a second
(.)	– micropause

:	– lengthening of a sound
◦	– low volume
£	– smiling voice / laughter
@	– voice quality
>huh<	– compressed or rush talk
<huh>	– slowed talk
.hhh	– hearable in-breath
hhh/HHH	– hearable out-breath
.	– pitch fall at the end of an intonation unit
?	– pitch rise at the end of an intonation unit
,	– level pitch at the end of an intonation unit
↓	– mid-turn fall in pitch
↑	– mid-turn rise in pitch
(-)	– inaudible word
+%	– embodied actions

References

- Bartsch, J. (2014). Beware the Frankenbite. *ThePowerEdit.com* [online]. Retrieved from <http://thepoweredit.com/beware-the-frankenbite>
- Couper-Kuhlen, E. (2012). Exploring affiliation in the reception of conversational complaint stories. In A. Peräkylä & M.-L. Sorjonen (Eds.), *Emotion in interaction* (pp. 113–146). Oxford, England: Oxford University Press.
- Curl, T. S. (2005). Practices in other-initiated repair resolution: The phonetic differentiation of ‘repetitions’. *Discourse Processes*, 39(1), 1–44.
- Curl, T. S., Local, J., & Walker, G. (2006). Repetition and the prosody–pragmatics interface. *Journal of Pragmatics*, 38, 1721–1751.
- Dingemanse, M. (2017). On the margins of language: Ideophones, interjections and dependencies in linguistic theory. In N. J. Enfield (Ed.), *Dependencies in language* (pp. 195–203). Berlin, Germany: Language Science Press.
- Dingemanse, M., Torreira, F., & Enfield, N. J. (2013). Is ‘‘huh?’’ a universal word? Conversational infrastructure and the convergent evolution of linguistic items. *PLoS ONE*, 8(11), e78273. doi:10.1371/journal.pone.0078273.
- Goffman, E. (1978). Response cries. *Language*, 54(4), 787–815.
- Goodwin, C. (1996). Transparent vision. In E. Ochs, E. A. Schegloff, & S. A. Thompson (Eds.) *Interaction and grammar* (pp. 370–404). Cambridge, England: Cambridge University Press.

- Hakulinen, A. (2001). On some uses of the discourse particle *kyl(lä)* in Finnish conversation. In M. Selting & E. Couper-Kuhlen (Eds.), *Studies in interactional linguistics* (pp. 171–198). Amsterdam, The Netherlands: John Benjamins.
- Heritage, J. (2011). Territories of knowledge, territories of experience. Empathic moments in interaction. In T. Stivers, L. Mondada, & J. Steensig (Eds.), *The morality of knowledge in conversation* (pp. 159–183). Cambridge, England: Cambridge University Press.
- Hoey, E. M. (2014). Sighing in interaction: Somatic, semiotic, and social. *Research on Language and Social Interaction*, 47(2), 175–200.
- Hutchby, I. (2006). *Media talk: Conversation analysis and the study of broadcasting*. Maidenhead, England: Open University Press.
- Jääskeläinen, A. (2013). *Todisteena äänen kuva. Suomen kielen imitatiivikonstruktiot* (Doctoral dissertation). Finnish, Finno-Ugrian and Scandinavian studies, University of Helsinki.
- Keevallik, L. (2013). Accomplishing continuity across sequences and encounters: no(h)-prefaced initiations in Estonian. *Journal of Pragmatics*, 57, 274–289.
- Keevallik, L. (2018). Sequence initiation or self-talk? Commenting on the surroundings while mucking out a sheep stable. *Research on Language and Social Interaction*, 51(3), 313–328.
- Keevallik, L., & Hakulinen, A. (2018). Epistemically reinforced *kyl(lä)/küll*-responses in Estonian and Finnish. Word order and social action. *Journal of Pragmatics*, 123, 121–138.
- Kidwell, M. (2009). “What happened?” An epistemics of before and after in “at-the-scene” police questioning. *Research on language and social interaction*, 42(1), 20–41.
- Legewie, N., & Nassauer, A. (2018). YouTube, Google, Facebook: 21st century online video research and research ethics. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 19(3). doi:10.17169/fqs-19.3.3130
- McIlvenny, P. (2015). The joy of biking together. Sharing everyday experiences of vélomobility. *Mobilities*, 10(1), 55–82.
- Norrick, N. R. (2007). Discussion article: Pragmatic markers, interjections and discourse. *Catalan Journal of Linguistics*, 6, 159–168.
- Ogden, R. (2006). Phonetics and social action in agreements and disagreements. *Journal of Pragmatics*, 38(10), 1752–1775.
- Pehkonen, S. (forthcoming/2020). Huh huh kannanottona fyysiseen rasitukseen. *Virittäjä*.
- Pomerantz, A. (1984). Agreeing and disagreeing with assessments. Some features of preferred/dispreferred turn shapes. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 57–101). Cambridge, England: Cambridge University Press.
- Reber, E. (2012). *Affectivity in interaction: Sound objects in English*. Amsterdam, The Netherlands: John Benjamins.
- Reynolds, E. (2016). *Incitement in the gym: Using pitch as a resource for action*. Philadelphia, PA: National Communication Association.
- Reynolds, E. (2017). Description of membership and enacting membership: Seeing-a-lift, being a team. *Journal of Pragmatics*, 118, 99–119.
- Robles, J. S. (2012). Troubles with assessments in gifting occasions. *Discourse Studies*, 14(6), 753–777. doi:10.1177/1461445612457490
- Schegloff, E. A. (1997). Practices and actions: Boundary cases of other-initiated repair. *Discourse Processes*, 23(3), 499–545.
- Sikveland, R. O. (2019). Failed summons: Phonetic features of persistence and intensification in crisis negotiation. *Journal of Pragmatics*, 150, 167–179.
- Sikveland, R. O., Skovholt, K., Stokoe, E., & Skarbø Solem, M. (2018, November). *Phonetic features of third turn receipts in Norwegian secondary school desk talk*. Paper presented at Fifth Nordic Interdisciplinary Conference on Discourse and Interaction, Aarhus, Denmark.

- Sorjonen, M.-L., & Vepsäläinen, H. (2016). The Finnish particle *no*. In P. Auer & Y. Maschler (Eds.), *Nu/nå. A family of discourse markers across the language of Europe and beyond* (pp. 243–280). Berlin, Germany: De Gruyter.
- Stevanovic, M. (2012). Prosodic salience and the emergence of new decisions: On approving responses to proposals in Finnish workplace interaction. *Journal of Pragmatics*, *44*, 843–862.
- Stevanovic, M., & Peräkylä, A. (2015). Experience sharing, emotional reciprocity, and turn-taking. *Frontiers in psychology*, *6*, 450. doi:10.3389/fpsyg.2015.00450
- Stivers, T. (2008). Stance, alignment, and affiliation during storytelling: When nodding is a token of affiliation. *Research on Language and Social Interaction*, *41*(1), 31–57.
- VISK=Hakulinen, A., Vilkuna, M., Korhonen, R., Koivisto, V., Heinonen, T. R., & Alho, I. (2004). *Iso suomen kielioppi*. Helsinki, Finland: Suomalaisen Kirjallisuuden Seura. <http://scripta.kotus.fi/visk> URN:ISBN:978-952-5446-35-7
- Walker, T., & Benjamin, T. (2017). Phonetic and sequential differences in other repetitions in repair initiation. *Research on Language and Social Interaction*, *50*(4), 330–347.
- Wharton, T. (2003). Interjections, language and the ‘showing’/‘saying’ continuum. *Pragmatics and Cognition*, *11*(1), 39–91.
- Wiggins, S. (2013). The social life of ‘eugh’: Disgust as a form of evaluation in family mealtimes. *British Journal of Social Psychology*, *52*(3), 489–509.
- Wilkinson, S., & Kitzinger, C. (2006). Surprise as an interactional achievement: Reaction tokens in conversation. *Social Psychology Quarterly*, *69*(2), 150–182.
- Wilson, D., & Wharton, T. (2006). Relevance and prosody. *Journal of Pragmatics*, *38*, 1559–1579.