

Clicks and percussives in English conversation

Richard Ogden

Department of Language & Linguistic Science, University of York

richard.ogden@york.ac.uk

Clicks are known to occur in English conversation, and have traditionally been assumed to convey affective meaning, generally negative. This is indeed the lay interpretation of clicks. In this paper, we build on the work of Wright (2011a, b), who shows that clicks are also used in the management of sequences of talk. Firstly, we consider similarities and differences in the production of clicks and percussives. We consider some distributional properties of clicks in one variety of English. Drawing on a wide range of conversational data from Britain and the USA, we show some of the functions of clicks and percussives in conversation, which include stance-taking and sequence management (projection of a turn constructional unit and word-search), and handling aspects of timing between turns. We also consider some of the visual behaviours that may accompany clicks. The meaning or function of clicks and percussives, it is argued, must be considered in a fuller interactional context.

1 Introduction

Although clicks are well known to occur in English conversation, there remain few studies which treat them anything more than anecdotally. In this paper, we survey some of the main forms and functions of clicks and their relations, percussives, from several collections of conversational data, and suggest possible future research.

Clicks in English are a worthwhile object of study for several reasons. Although they do not occur in lexical items in English, they are widely assumed, both by linguists and lay people, to have a communicative function. A study of clicks is therefore an exercise in excavating non-lexical function and meaning from everyday talk. This paper is a contribution to a growing tradition that approaches the problem by combining phonetic analysis with conversation analytic techniques as a way to demonstrate how participants themselves understand and orient to one another's displayed meanings (see e.g. Ogden 2012a). Clicks have been shown to be involved in managing conversation (Wright 2011a, b), and are frequently implicated in displays of stance or affect (Reber 2012). While there are theories which aim to connect the form and function of prosodic features such as loudness or intonation contours (see e.g. Gussenhoven 2004 for an account using biological codes), such theories typically consider ways of modifying lexical material, and have little or nothing to say about punctual – i.e. instantaneous, transient – phonetic events like clicks.

Clicks are embedded within stretches of speech, and are 'in' speech but not 'of' it. They are sounds made in the vocal tract alongside speech, not as part of the lexical content of the language, but clearly as a resource for making meaning. The same claim could be made for many co-speech sounds, such as laughter (see e.g. Kohler 2007 for a description of varieties

of co-speech laughter, and e.g. Jefferson 1984, Potter & Hepburn 2010 on the positioning of laughter in conversation), sighing, sobbing, or audible inhalation and exhalation. As Laver (1994) points out, communication consists of several semiotic systems. In developing an understanding of the functions of clicks, we are exploring a sound that is describable in phonetic terms but which is considered normally to be at the margins of language itself, along with other semiotic systems such as gesture, facial expression, and other vocal and bodily movements and sounds. A study of clicks in a language where they have non-lexical functions, then, necessarily involves thinking about where the boundaries of language lie.

Gil (2011) claims that cross-linguistically clicks are typically paralinguistic in function because they are ‘not integrated into morphological and syntactic structures, [and] convey a restricted range of meanings, some of which are associated with the expression of emotions’ (p. 1). He claims that in some languages – Hebrew, Avar, Somali, Fula and Malayalam, among others – clicks are able to serve as a response to yes/no questions. In both cases, clicks resemble linguistic signs in that they are arbitrary and conventionalised, even though they are not lexical.

For English, the primary function of clicks is generally held to be stance-taking. Clicks are mostly – it is claimed – negatively valenced. Wright (2011a: 208) provides a summary: ‘disapproval (Ladefoged 1982: 124; Crystal 1987: 126), annoyance (Abercrombie 1967: 31; Ball 1989: 10), irritation (Gimson 1970: 34), exasperation (Laver 1994: 175), impatience (Laver 1994: 175), regret (Clark & Yallop 1990: 18), sympathy (Gimson 1970: 34) and encouragement (Gimson 1970: 34; Abercrombie 1967: 31; Ladefoged 1982: 124; Laver 1994: 177)’. Most of these labels come from anecdotal observation, and few are exemplified using empirical data. Reber (2012) claims, from an examination of natural conversations, that stand-alone clicks as responses to complaints about third parties or external events display a weak affiliative stance as opposed to strong moral indignation, which is done through expanded turns at talk. Ward (2006) subsumes the meanings of clicks under ‘personal dissatisfaction’, perhaps with something another person has said, the current topic, or (in the case of word search) with one’s own performance. As we shall see, the meaning of clicks is more nuanced than this characterisation indicates; and we shall also argue against the essentially mono-systemic account of clicks which Ward assumes.

The lay interpretation of clicks in English-speaking cultures agrees that clicks display a negative stance or affect. This is conventionalised through the word ‘tutting’, which refers not just to the sound of a click, but also to the negative stance that clicking conveys. Example (1), taken from a radio phone-in in Britain, illustrates this well. (Transcription conventions are explained in Section 3 below.)

(1) vegtalk 19.12.03.03:33

```
01  E      there was aNOther british fAMILY `THERE,
02          who were? h↓ rEAlly em`BARrassed;
03          cos I gOt Up and made a `SPEctacle of mysElf;
04          and thEy were all LOOKing [.hhh [ahaha
05  P →          [!      [!
06  E      <<laugh> dOwn at the[ir>
07  P →          [ `TUTting and
08          pretEnding to dis`OWN you.=
09          =I knOw [how it IS; Ellie yEs yEs]
10  E          [<<laugh>----->]
```

In lines 1–4, the caller, Ellie, describes an embarrassing situation resulting from her own inappropriate behaviour (line 3). In line 5, the presenter makes two click sounds, [! !],

displaying an understanding of the situation as one which elicits disapproval. In line 7, he glosses this expression of disapproval as *tutting*. The verb form in line 8, along with his display of sympathy towards Ellie in line 9, retrospectively cast this disapproval as originating from the *really embarrassed* family watching Ellie as she *made a spectacle* of herself. *Tutting*, then, is a recognisable activity, occasioned by socially sanctionable behaviour.

Stance-taking is not the only use of clicks. Wright (2005, 2007, 2011a, b) shows that in normal conversation, clicks can also handle aspects of sequence management. They commonly occur at boundaries between sequences, such as those found at the end of a phone call, where the speakers have to close down a current topic before initiating the closing of the call. This can be seen in Example (2): lines 1–5 close a prior sequence, and line 7, [!] *okay, well then*, initiates the closing of the call with arrangements for next seeing each other (see e.g. Schegloff & Sacks 1973, Button 1987 on call-closing practices in English-speaking cultures).

(2) ESF/O41 SBL:2:2:3:R:60–61

01 Zoe wElI it was [`FUN `CLAIRe;
 02 Claire [yEA:[h
 03 Zoe [mm [(*
 04 Claire [?I enjOYed every `MINute of [it
 05 Zoe [yAp`
 06 (0.4)
 07 Claire → ! ↑o`KAY, well ↑-thEn we'll `sEE you: (.)
 08 `SATurd[ay.
 09 Zoe [↑-SATurday `NIGHT.
 10 Claire <<h> sEven `THIRty?>
 11 Zoe `yAp`.

Wright (2005) also observes that clicks are commonly produced during word searches. We will extend Wright's work on word searches in this study, but will not provide more instances of clicks which index a new sequence.

The structure of the rest of this paper is as follows: First, we consider the phonetics of clicks and their close relatives, percussives. Next, we present some simple quantitative data on the distribution of clicks and percussives in a sample of English conversational data; and then we consider more closely the forms and functions of clicks in some of the sequences of talk where they are found. We will also look at aspects of timing of clicks and accompanying visible actions. We will argue, on the basis of the data presented, that clicks in English are implicated in three distinct systems with distinct functions but partially overlapping phonetic (and kinetic) exponents: (i) marking incipient speakership; (ii) handling aspects of sequence management; and (iii) as part of a display of stance.

2 Phonetic overview

Clicks are ingressive, velarically initiated suction stops (Catford 2001). They can be co-produced with various articulatory accompaniments, such as nasal airflow, voicing, and phonation types (Ladefoged & Traill 1994). In English, most such accompaniments are rare. There is occasional aspiration, but clicks are often followed by one of a large number of response tokens such as *oh*, *ah*, *aw*, *oo*, which – in contrast to the click itself – may be articulated in a wide range of ways, including phonatory setting, loudness, duration, voice quality, vowel quality, and glottal stop initiation. These are not the main focus of this paper, but, as will be clear, these tokens and their modifications are crucial in the display of affect.

The International Phonetic Alphabet classifies clicks by place of articulation. Nonetheless, Ladefoged & Maddieson (1996) argue that the articulation of clicks is complex, and place of articulation may not be the most satisfactory or explanatory way to describe them: the speed of release, the amount of tongue tension, suction, and tongue shape are all important aspects of the production of clicks with complex acoustic correlates.

Given the difficulties mentioned, we represent lingual English clicks in transcription with the symbols [!] and [!]. [!] is used to refer to a click articulated with the tongue tip, with central release which is generally slow and affricated. The closure may be dental or alveolar (Wright 2011a: 215), but the release nearly always has affrication, which implies a relatively slow release and a relatively large contact between the tip/blade of the tongue and the roof of the mouth. This is the click generally represented in English texts as ‘tsk-tsk’ or ‘tut-tut’. This click is ubiquitous in English. The second kind of click that is common in English, but considerably less common than [!], is a lateral one, [!]. This click also has closure with the tongue tip which impressionistically ranges from dental to palatal (and is retroflex), but the release is lateral. Compared to [!], the lateral click has a deeper, hollower ring, and a sharper release. This click is comparatively rare in the data in this paper.

Closely related to clicks in English are sounds made as the articulators separate and the speaker prepares to speak (Scobbie, Schaeffler & Mennen 2011). We call these sounds PERCUSSIVES, following Pike (1943: 103–105):

Percussors differ from initiators in several ways: in opening and closing they move perpendicularly to the entrance of the air chamber . . . ; they produce no directional air current, but merely a disturbance that starts sound waves which are modified by certain cavity resonators; they manifest their releasing or approaching percussive timbre only at the moments of the opening and closing of some passage . . . (Pike 1943: 103)

Articulations which Pike specifically names as percussives are the opening and closing of the lips, the tongue making closure at the alveolar ridge, the velum closing, the vocal folds making a glottal closure, and the sublaminal percussive of the ‘cluck click’ as in ExtIPA’s [!j]. For Pike, then, percussives arise from either separation of articulators or the striking together of articulators. Percussives are not audible unless the movement is ‘sharp, forceful and rapid’ (Pike 1943: 103), and when the resonating cavity is closed. Pike describes the initial burst of plosives as percussive, but says that it is not convenient to separate them from the rest of the plosive. In this paper, then, we will use ‘percussive’ to describe sounds arising from the audible separation of articulators which are most likely to be produced as pulmonic ingressive plosives; but the term ‘plosive’ suggests a deliberate articulation of a sound found in English words, and it is not clear that percussives are always deliberate. They seem, rather, to be often ‘organic’ or ‘vegetative’ sounds which arise as participants position themselves as an incipient speaker audibly gearing up to speak. These articulations serve as an audible and often visible signal that an individual is preparing to speak.

The Extensions to the IPA offer the symbols [w̥] and [ɸ̥] for bilabial and bidental percussives respectively. These symbols refer to ‘the striking together of two rigid or semi-rigid articulators’ (IPA 1999: 187), a usage of the term ‘percussive’ which is more restrictive than Pike’s. The sounds discussed here are generated by the separation of articulators, which makes [w̥] and [ɸ̥] unsuitable. We shall therefore use instead a combination of the symbol for pulmonic plosives, to represent the location of the release, and the symbol for ingressive airflow, [↓]. Audible in-breaths are a regular feature before and after both percussives and clicks; so it is pragmatic and linguistically informative to transcribe all percussives in a consistent manner, as [p↓, p̄t↓, p↓h↓], etc. The in-breath is most audible after the release of the percussive or click, but quite often inhalation is audible and visible on spectrograms.

Acoustically, clicks are expected to show two transients corresponding to the anterior and velar releases. An example of this is shown in Figure 1, which contains two transients, marked

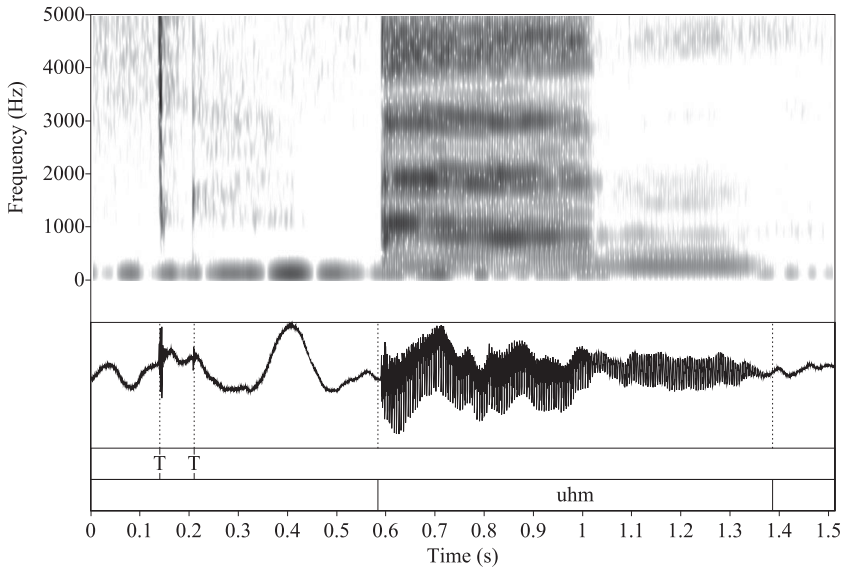


Figure 1 [! h↓ ʔɜ::m]. (Speaker from Salford; first part of a response to a question.) The two transients (marked 'T') correspond to the release of the anterior and posterior closures for the click [!].

'T'. The first one corresponds to the anterior release, the second to the dorso-velar release. In general, clicks are louder than percussives. However, they are variable and the transients that must occur are often not easily seen through the high-amplitude friction noise which often accompanies clicks.

Example (3) illustrates percussives in initial position in a Turn Constructional Unit (TCU), the basic unit out of which turns at talk are composed (Sack, Schegloff & Jefferson 1974; see also Section 5 below).

(3) Salford Si&Sh 709–717 drinking

- 01 Sh it used to be gOod on the `HEIght, `drInking,=
 02 Si =nOw it's just gOne (.)
 03 the yOUng uns have `SPOILT it.
 04 (0.4)
 05 Sh → t↓k↓ h↓ it's `RUBbish;
 06 (0.5)
 07 Sh so we dOn't go `OUT there any mOre.

Before Sheila produces *it's rubbish*, she audibly breathes in and there are a number of transients in the speech signal. Auditory analysis suggests the tongue coming away from the alveolar ridge and velum almost simultaneously, hence the transcription [t↓k↓ h↓].

Figure 2 shows a spectrogram of the percussives in Example (3). The first one is alveolar, while the other two transients are velar and show a typical velar burst with multiple releases. It can be seen that percussives can involve the release of multiple closures, since in the rest or closed position, the tongue may make contact along its whole surface and length. It is not difficult to see how at least some clicks may arise from such ingressive percussives: it only requires an anterior contact to be released before a velar contact, and

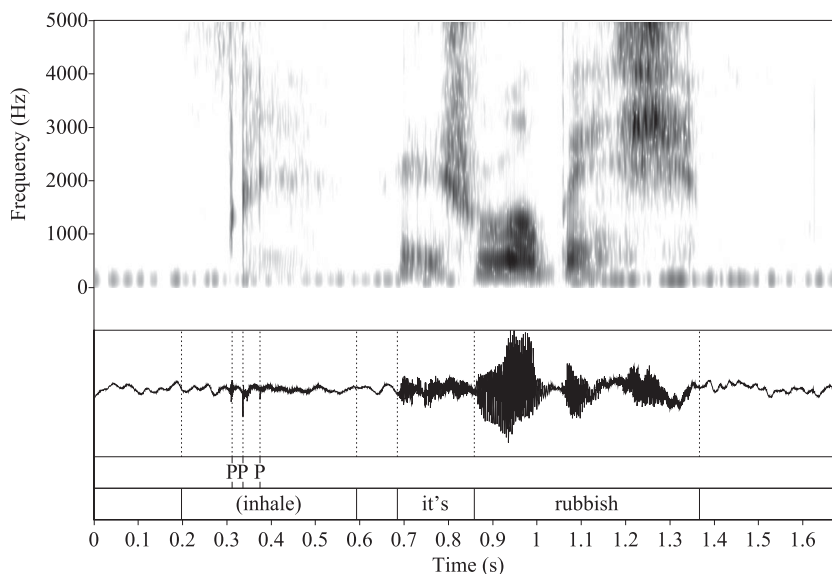


Figure 2 Example (3), line 5: *it's rubbish*. The transients around 0.2 s are percussives (marked P in the waveform and visible in the spectrogram), accompanied by an inhalation till about 0.4 s.

some amount of suction on the tongue to produce the supralingual rarefaction required for a velaric ingressive suction stop. The amount of tongue pressure will control the amount of rarefaction, and thereby also the loudness and salience of the click. This means that some clicks are quiet, especially in turn-initial position, and some clicks and percussives are very hard to distinguish auditorily. As we will see, in at least some structural positions, clicks and percussives are in free variation, and it may be that the more communicative – and auditorily much more salient – clicks which handle sequence management and displays of stance have evolved from the quieter turn-initiating clicks which alternate with percussives. We will not in this paper explore this much further, because acoustic data do not easily allow it.

3 Data

The data for this paper come from a wide range of sources, all of them collected for other purposes than the current one. The sources of the examples in this paper are identified by a mnemonic in the header line; these are described in brief in the Appendix. All the data are conversational and include material from radio phone-ins, phone calls and sociolinguistic interviews.

The extracts presented here are transcribed using a version of the GAT conventions (Gesprächanalytisches Transkriptionssystem) for the transcription of conversation (Selting et al. 2009), which mark features of talk using adaptations of orthography without losing readability. GAT uses the symbol ! to mark strong emphasis; this has been replaced with underlining to avoid confusion with the click [!]. Clicks and percussives are transcribed using IPA conventions, with some modifications as described in Section 2 above. Since square brackets are used for marking overlapping talk, any segmental IPA symbols in the transcriptions are presented in italics, to avoid confusion.

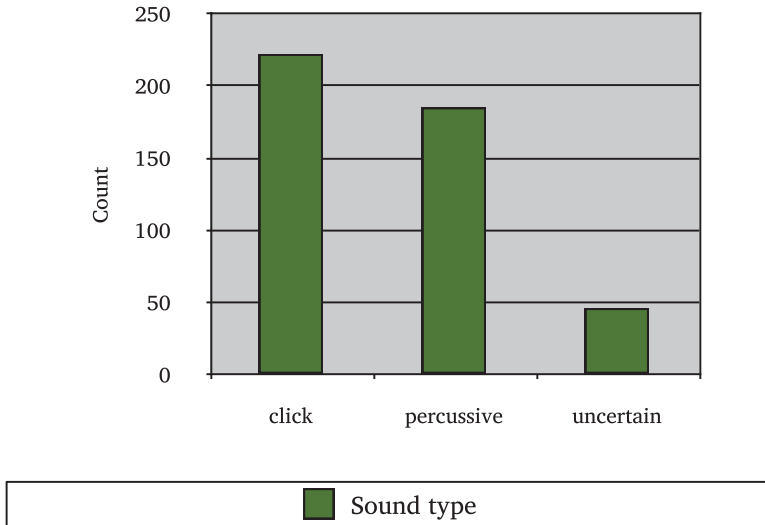


Figure 3 Clicks and percussives (including uncertain cases) in a corpus of spoken data from York.

4 Patterns of distribution

To give some idea of how common clicks and percussives are, we report a survey of one subset of the data, the York section of a set of sociolinguistic interviews collected as part of a project on Northern Englishes.

The data consist of four male pairs and three female pairs. All speakers were in the age range 18–22 at the time of recording. All but two were in tertiary education, and all were born and brought up in York. There are about 40 minutes of material for each pair, i.e. about 280 minutes in total. Clicks and percussives were identified by three different transcribers, who labelled separately and then compared labels and agreed on the classification of each sound using acoustic and auditory evidence. Where sounds were quiet, transcribers could not confidently distinguish between clicks and percussives, so we report an ‘uncertain’ category as well.

Clicks and percussives were approximately equally numerous (Figure 3), and there was a substantial number of cases where transcribers could not reach agreement on the classification. In all, there were 222 clicks which the transcribers were certain of, which means that in five minutes of talk, on average, about four clicks occur. A similar frequency was found for Finnish (Pohjola 2011).

If clicks are considered by position in the turn, the majority (128, 57%) are TCU-initial; 90 (41%) are TCU-medial, and a very small number (4, 2%) are TCU-final; but it is likely that these clicks project a new TCU which is abandoned. Clicks can be followed by a range of vocalisations, such as *uhm*, *mm*, particles such as *oh*, *ah*, *aw*, or linguistic items such as words. Pauses are also common before and after clicks.

When the distribution is considered by speaker (Figure 4), it is obvious that some speakers produced a very large number of clicks: three speakers (Anna, Emma and Cathy) account for almost half the clicks in the whole collection. This is an interesting finding, but it is also hard to interpret. Are some speakers habitual clickers? If so, why? Is it the case that these three speakers (but not the others) produced a large number of turns at talk where clicking was relevant? Is it possible that some speakers produce a lot of clicks but not many percussives, and vice versa? If so, this would suggest that clicks and percussives are indeed in free variation, and the choice is idiosyncratic. Considering the numbers by speaker does

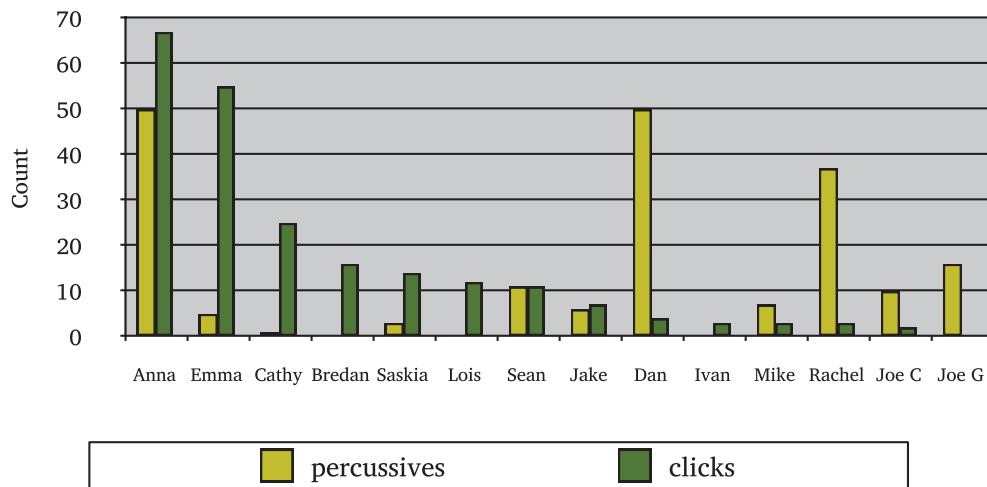


Figure 4 Clicks and percussives by speaker, arranged in descending order of the number of clicks produced. The patterns of usage look idiosyncratic.

not solve the problem: Anna produces many clicks and percussives, Emma produces many clicks but not percussives, while Dan and Rachel produce many percussives but not clicks. So it is not clear that clicks and percussives are in free variation and one kind of sound can freely be substituted for the other.

Of 222 clicks identified, only six were clearly identified as nasalised; another six were identified as possibly nasal, and the transcribers could not reach agreement. Nasalised clicks such as [ŋ̥] are therefore very rare in this dataset, and comprise no more than about 5% of all clicks. No instances of voiceless nasal clicks were identified.

If clicks are grouped by speaker gender, then females overall click much more than males: females produced 176 of all the clicks (= 79% of all clicks). On average, each female speaker produced 29.3 clicks in 40 minutes, i.e. approximately one click every 1 minute and 20 seconds, while each male speaker produced 5.75 clicks in 40 minutes, i.e. approximately one click every seven minutes. The three biggest clickers in the collection are females, and they produce 147 (66%) of all the clicks, which indicates that clicks are not JUST a matter of gender. It is hard to know how to explain this finding. It may be that these speakers produce more of the kinds of talk where clicks are relevant. It may be that the clicks have a social-indexical function. To provide an explanation, a better understanding is needed of how clicks are distributed in English conversation, and what functions they have. The next section addresses this issue.

5 Turn-initial clicks and percussives

In this paper, we assume that speakers' turns at talk are organised as described by Sacks et al. (1974). To simplify the details of that paper: The basic unit of organisation is a turn, which is the time during which a participant is speaking. Turns consist of one or more TCUs (Turn Constructional Units), which are of variable size, but contain lexical material. TCUs usually contain places, Transition Relevance Places (TRP), where a current listener may legitimately come in with a response: this is best thought of as a space rather than a point. Ford &

Thompson (1996), among others, show how TRPs are aligned with syntactic, pragmatic and prosodic structures. In linguistic terms, TCUs therefore are best thought of as constructions describable on several levels simultaneously.

The commonest location for clicks is the pre-turn position: that is, before any lexical content in the turn, and before any response tokens such as *oh* or *aw*. Many such initial clicks are ‘plain’, in two senses: they are voiceless and non-nasal, and have only an in-breath as an accompaniment (though not necessarily), and they seem to do little more than mark the initiation of a turn at talk.

Examples (4) and (5) illustrate ‘plain’, turn-initial clicks. Example (4) is taken from a radio show, and P is the presenter.

(4) ell lists sum04 cheese

01 P ↑whAt's on Offer at the mArket then to`DAY.
 02 I → ! h↓ ↑-wE:l1- we've gOt lots of -CHEEses==
 03 =we've got h↓ -PANcAkes==
 04 =very h↓ tradItional cOntinEntal -DISHes-

The question at line 1 gets a reply at line 2. This reply is initiated by *well*, which marks it as not a straightforward answer (Schegloff & Lerner 2009), and this non-straightforwardness comes in the form of a list. It would be socially odd for the click to mark a negative stance in this context, and there is no other material in the turn that displays a negative stance. A reasonable interpretation of clicks like these is that they are an audible signal of incipient speakership, an interpretation that is consistent with the subsequent in-breath. Speculatively, in multi-party talk, it may be that such clicks can be used to identify oneself as a next speaker, and to secure the attention of others. As relatively loud transient sounds, clicks may be well adapted to this purpose.

Example (5) resembles (4) in that it is a responsive turn. However it is sequentially more complex because the turn starts with a new action (a greeting), and is the start of a longer activity, of responding to Imogen's point. In this respect it is similar to Wright's (2011b) New Sequence Indexing clicks.

(5) klm lists sum04 housing

01 I ((complains about new houses to be built in M-))
 02 M and you're un`HAPpy abOUt thAt.
 03 I yEs;=and so's a lOt of pEOple in `M-.
 04 M `THANK you?
 05 COUncillor `P-?
 06 P → ! h↓ `O:h;=↑hel'LO `Imogen?
 07 the `ISSue `Is, as you knOW, is the GOvernment...

Here, a caller to a radio show, Imogen, has been complaining about new houses being built in her town. The moderator selects Councillor P as the next speaker in line 5. P's initial action on starting her turn is to greet Imogen, followed by a reply to Imogen's earlier comment. P's turn starts with a click, followed by an in-breath and the turn-initial particle *oh*. As in (4), it is hard to make a case for this click marking a stance; more likely is that the click is an audible sound of ‘gearing up to speak’ (see Scobbie et al. 2011), as the speaker separates her articulators and takes a breath in.

Clicks in these structural positions seem to be in free variation with percussives, as seen in the arrowed turns in Examples (6) and (7), which are initiated with percussives.

6 Stance-taking

As we have seen, clicks may be used as a device for projecting up-coming talk: they occur turn-initially and are followed by what has been projected in a prior turn. We look now at clicks that are closer to the lay interpretation, i.e. clicks which are implicated in displaying a stance. We shall see that the form of such clicks is in general no different from those of turn-initiating clicks, but their sequential placement provides them with an affect-laden interpretation. Percussives seem not to occur in these contexts. Example (9) provides a good example of a click as part of a conveyance of negative stance: a canonical ‘tutting’ click.

(9) CallHome 4861.60

01 M Obviously I didn't do a gOOD enough
 02 j[Ob of `RAIS[ing you.]
 03 D → [! h↓ [<ah -s]t0:p -thA:t->

Mom has been discussing at length with her daughter Debbie why Debbie has not yet found a suitable prospective husband, and has not settled on a career. At lines 1–2, Mom blames herself for what she perceives as her daughter's failings. Debbie's response consists of a click followed by an in-breath, and then an imperative which rejects Mom's self-blame and seeks to close it down. The form of the click is the same as in earlier examples: a plain voiceless click accompanied by an in-breath. We saw no evidence in other examples that such clicks were displaying a stance: they mark only that the speaker is getting ready to talk. If we consider the placement of Debbie's click relative to Mom's talk, it comes much earlier than the other [plain click] + [in-breath] constructions. It is placed at a point in time where Mom's self-blame becomes unambiguously recognisable as such (*obviously I didn't do a good enough j-*). Thus by virtue of its timing relative to the prior turn, this click projects an immediate response by Debbie to Mom's self-blame; and indeed her resistance in *oh stop that* is located exactly at a TRP, just after the main accent of Mom's turn, in *raising* (see Wells & MacFarlane 1998). Thus the click projects a turn dealing with Mom's self-blame by its placement; but the form of the click does not, *per se*, express resistance to Mom. Any sense of resistance arises from the click's timing relative to Mom's on-going turn, and the accompanying actions expressed through *oh stop that* which together deliver a stance towards Mom's turn.

Example (10), discussed in more detail in Drew (1998), illustrates another click expressing disapproval.

(10) Holt C38:4

01 L and hE came Up to me and he sAId
 02 "Oh hello LEsley, †stIll trYing to bUY sOmething for "NOthing?"
 03 J → ! q↓ :: [^OOH.
 04 L [↑^OOH.
 05 J Isn't [he
 06 L [↑whAt do you `SAY.↑
 07 J Oh Isn't he ^DREAAdful.

Lesley has been telling Joyce a story about a mutual acquaintance, *Mr R*, at a sale in her local vicarage. Lesley launched her story by saying *I'm broiling about something*. The punchline of her story is at lines 1–2. Mr R's reported words in line 2 are the thing which she is *broiling* about; and Joyce's response at line 3 is her appreciation of Lesley's story. Joyce displays her shared outrage at Mr R's behaviour with a click followed by a long voiceless ingressive sound made with an open vocal tract, [! q↓::]: it is close to a 'gasp'. Joyce's turn thus does two things:

it simultaneously displays an understanding of the point of Lesley's story, and affiliates with Lesley's own stance towards Mr R's comment to her. It is followed by further talk in which Lesley and Joyce share their outrage at Mr R.

The TCU which is initiated with the click conveys several actions at once: it displays 'outrage' at the complainable behaviour reported in Lesley's story, which is projected right at its beginning and in doing so, the TCU affiliates with Lesley in her complaint. In this case again, the click is understandable through its context (after the punchline of a story, and the actions that it makes relevant from the recipient of the story); but there is nothing in the form of the click itself which presents it as a click designed to do these things.

In Example (11), the click is again turn-initial and accompanied by a response particle, this time *ah*, which is produced with breathy voice quality, and followed by an assessment, *that's lovely*. As in (10), the click is in a turn which displays an understanding of the kind of story that has just been told; in this case, a romantic one, where a young woman has finally got together with a man who has been very persistent. Kate's appreciation of the story is positive.

(11) nrb reluctant lover

01 J he wOUldn't stOp Asking her `OUT.=
 02 =he used to rInG her like thrEE times a `DAY;=
 03 =and she'd go ``NO; `NO;“
 04 Or she'd sAY yEs and nOt turn `Up;
 05 and thEn she jUst complEtely `↓FELL ↑for him-
 06 K → !^h (.) ↑`ʔɑ: [thAt's `lOvely.
 07 J [to[gEther

Examples (10) and (11) display two different kinds of story which are differently appreciated, but the appreciation of the story is initiated with a click. The click does not occur by itself though. Although the click marks the onset of the appreciation of the story, the KIND of appreciation that the recipient of the story is expected to display is projected at the launch of the story, and through the telling of the story itself. The click is accompanied by a response token which is fitted to the story. The [click] + [response token] are followed by further lexical material which displays verbally an appreciation of the story. Thus the 'meaning' of the click is part of a more complex set of interactional practices for story-telling.

In Example (12), there is a rare case of a nasalised click. Kate has dyed her hair red, and Jade is complimenting her on the result. In line 4, she makes a positive assessment of Kate's hair, and Kate responds in overlap, but at a TRP. Like the clicks that we saw in stories, it is accompanied by a response token ([ʔɑ:], with intonation falling to low), and then verbal material which explicates the action of the turn, in this case thanks.

(12) nrb red hair

01 J ʔs vERy RED; but it's vERy `NI[Ce.
 02 K [it Is vERy RED
 03 [Isn't it.
 04 K [lOOks in rEAlly good con`DITi[on.
 05 J → [ŋ~!`ɑ: [thAnk yOU]
 06 K [(*) lOvely and] shIny

In summary, clicks that display a stance are often accompanied by other material which conveys a stance. This is expectable. Response tokens such as *oh*, *ah*, *aw*, etc., are vocalic and thus amenable to a range of phonetic modifications which serve to mark the kind of stance they display (Norrick 2012). Clicks, on the other hand, are much more difficult to modify:

they are punctual, they are usually voiceless, and modifications to them are more difficult to produce and to perceive. A reasonable conclusion is that clicks do not by themselves display a stance, but are part of a display of a stance which is handled and negotiated interactionally over a longer stretch of talk, e.g. through the projection of the kind of appreciation that is due, the use of response tokens, and the inclusion of overt lexical material which provides an assessment or displays gratitude. This is compatible with recent studies which show that many stance-related actions in conversation are not instantaneous, but negotiated over longer sequences of talk (e.g. Wilkinson & Kitzinger 2006; Local & Walker 2008; Ogden 2010, 2012b). The task for the producer of the click is to produce a click at the right place in a sequence, and at the right time. However, we have been unable to show that the particular form of the click is consequential: modifications to place of articulation or accompaniments to the click articulation do not seem to be used to modify what is displayed through the click. Rather, if clicks have any meaning, this comes about through other features of the collocations in which they occur. A possible exception are nasal clicks and lateral clicks, which feature only rarely in the data; further study is needed to show empirically whether these kinds of clicks have particular functions.

7 Kinetics

We now consider some aspects of the kinetic production of clicks and percussives. So far we have considered only the auditory/acoustic aspects of clicks, but there are frequently other accompanying behaviours involved in the production of both clicks and percussives.

One physical action which may accompany clicks is swallowing. In Example (13), there is a word search which contains the particle *uh*, a pause, and a click: it shares many similarities with Example (8) above.

(13) RCE25 Bench 00.57 library card

```
01  L          and thEn I went Over to the LIbrary to-uhm (0.6/swallow)
02    →      !~0 uh SEE abOUt thAt-uh
03          <<all, p> I cAN't remEmber the nAmE of the cARd now> but to
04          to use Other univErsit[y lIbraries.
05  R                                [oh a (scOnal) card.
```

A transcript based on auditory data alone at line 1 shows a pause of 0.6 seconds, but inspection of the video shows that L is swallowing during that part. The click is the culmination of the swallow, where L opens her vocal tract. Swallowing involves a movement of the tongue and pharyngeal muscles which produce a downward movement of the bolus (which in this case is likely to be saliva). This downward movement, accompanied by closure of the surface of the tongue against the roof of the mouth, and closure of the lips, is conducive to velaric ingressive airflow as the vocal tract is opened again.

Figures 5a–c are all taken in the 0.6 second silence between *to-uhm* and the click. Figure 5a is the bilabial closure at the end of [m]. Note that it is not merely bilabial closure involving the surface of the lips, but involves contact of the whole of the outer surface of the lips; this is a very extreme closure as compared to that normally used for speech. Figure 5b is taken at the point where the speaker swallows. In the video, the tendons in the neck are seen to become more prominent as the swallow occurs, and the cheeks are also drawn in. As can be seen, the lips are starting to be released, although there is still a tight closure. Figure 5c is taken at the point of release, where the click is made. The ‘lipsmack’ at the resolution of the word search is therefore not just audible; it is the culmination of a bigger physical process which is also visible from the speaker’s face and neck.

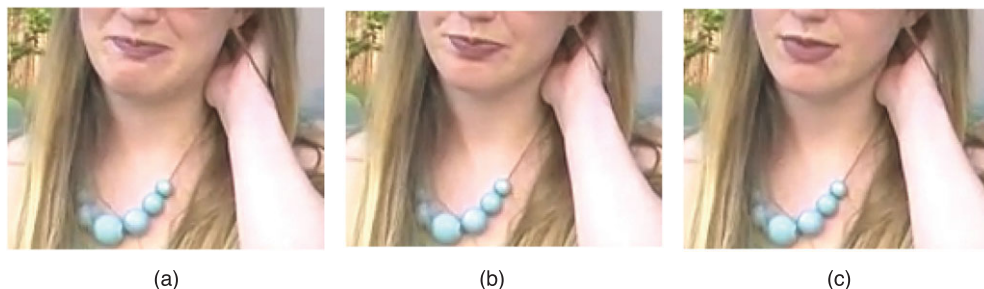


Figure 5 Example (13), lines 3–4: (a) bilabial closure; (b) mid closure, at the peak of the swallowing; (c) just before release into *to-uhm* (0.6) [1].

The next example of a kinetic event co-occurring with a click is taken from a family dinner in the United States in the 1960s. In previous examples, we saw that clicks that present a stance are typically accompanied by turn-initial particles and followed immediately by lexical material which expands on the action of the turn with the click, or of the [click] + [response token]. We turn now to a click which does occur by itself. In this example, Virginia, a teenager who is not yet old enough to work, is pestering her mother about getting a new dress (lines 1–2, 4). That this is not the first time she has done this is clear from Mom's turn in lines 5–6. Mom resists Virginia's request for a dress, to which Virginia responds at various points. Virginia's *uhhh!* at line 8 is accompanied by a head turn away from Mom.

(14) Virginia p.3 01:27

Context: Virginia (V) and Mom (M) at a family dinner table. (Some overlapping talk from Virginia's sister Prudence at line 13 has been removed for clarity.)

01 V cAn I plEAsE get that -DRE:SS-; -plEAsE MOM-
 02 lEmme g[Et that-
 03 M [↑DRESS;
 04 V <<all> you knOW that [one->
 05 M [↑OH virGINia,
 06 we've bEEn through thIs be'FORE,
 07 yOU've got enOUGH `sUmma d[rEsses now,
 08 V [<<br, p> `↑?UH;>
 09 M I thInk you just wAIt and gEt some OF the new `FALL stuff
 10 when it comes <<cr> i:n.>
 11 V → !
 12 (0.5)
 13 M if you sAved your- if you sAved your al`LOWan[ce;
 14 M (if you) sAved yer aLLOWance; an:' um: you could `GET
 15 M these little Extr[a things.
 16 V [a(h)l`LO: :Wan(h) ce?
 17 I O(h)nly g(h)Et fI(h)ve d(h)Ollars a WEEK;=
 18 =that's ri`D(h)I(h)c(h)ul(h)ous.

At line 11, Virginia produces a click. It is accompanied by a head-turn away from Mom, with Virginia shifting her gaze away, and iconically disengaging from her mother (Figure 6) (see Goodwin 1981, 1986).



Figure 6 Upper panel: Mom (left) and Virginia (right) as Mom says *fall stuff* (Example (14), line 9). They share mutual gaze. Lower panel: Mom (left) and Virginia (right) as Virginia produces her click (Example (14), line 11). Virginia and Mom are visibly disengaged from one another.

Virginia's click can be understood as part of a display of resistance to Mom on a number of grounds. Firstly, there are other, earlier signs of her disaffiliation with her mother's stance, and indeed signs of her mother's resistance to her: her request for the dress in line 1 is immediately followed by a repeated form *lemme get that-*. Her mother's response at lines 5–6 rejects the request – not for the first time – and in lines 7 and 9, Mom presents reasons why she does not go along with Virginia's request. Virginia in the meantime makes a whimpering noise (line 8), and turns away from her mother. The click then is located in an extended sequence where the two parties have a stand-off: on completion of Mom's accounts for why she will not grant the request, Virginia displays her response. The click is therefore but one part of how Virginia displays her stance towards Mom, both in the turn itself and in the sequence more generally. Clearly, Virginia's click is implicated in displaying a stance towards her mother, but it does not in and of itself display these things. While the click is the only audible material in the turn, it is not the only physical action that Virginia produces at this moment: Virginia's turn delivers an action through a blend of vocal and kinetic activities.

8 Clicks as metronomes

The final set of clicks we consider are those which are timed isochronously with other events in the speech signal. Clicks work uniquely well for this: as they are transient and relatively loud, they serve well as temporal markers in speech. For this reason, we claim that these clicks function as 'metronomes'; alongside any other functions they might have, they handle the intersubjective timing of participants' turns relative to one another.

The first case is a click which is part of a display of sympathy, in Example (15).

(15) SW/CK Sym #63

```
01  A      it's jUst kinda ^DULL;
02          GOd what a `m:Iserable `mIserable [(.)] [^WEEkend.
03  B                      [ ! ] [ `ɑ : : : :
04  B      thAt's a `SHA:me.
```

In this extract, A is complaining about the *dull* and *miserable weekend* she has just spent. In line 2, she repeats the word *miserable*, then there is a gap in her talk, before she produces *weekend*. In this gap, B comes in and produces a click, which is located just between the second *miserable* and *weekend* (see Figure 7). In overlap with *weekend*, B starts a turn which displays sympathy with A: it has the format [click] + [response token aw] + [verbal expression of sympathy]. This makes it structurally similar to other cases considered earlier. These two turns are produced with isochronous beats. In her turn, A produces words aligned with the beats (on *God*, *miserable*, *miserable* and *weekend*) with a 'beat of silence' between *miserable* and *weekend*. Figure 7 shows a labelled waveform with three tiers: A's words, B's words, and beats (marked b), as produced automatically using a program which places beats half way up amplitude contours in the speech signal, as in Cummins & Port (1998).¹ As can be seen, an isochronous rhythm holds across the two turns, with B's click produced on the fourth beat, and then three further beats on *aw*, *that's* and *shame*. The temporal placement of B's click thus demonstrates an orientation to the rhythm established by

¹ Code available from <http://cspeech.ucd.ie/~fred/beatExtraction.php>. The code places beats slightly before the expected location in the case of transients such as for clicks; the beats are therefore shifted leftwards by a few milliseconds. The isochronous beats are nonetheless visible by this method, which has the advantage of being consistent and placed by a computational algorithm.

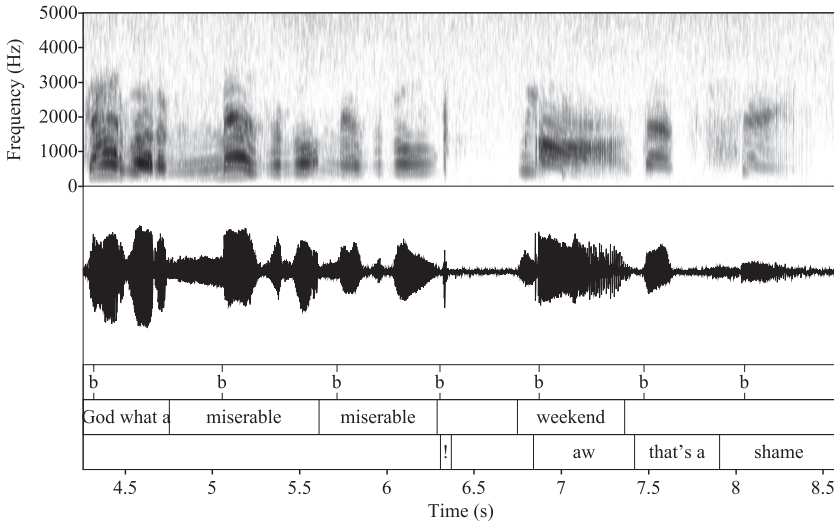


Figure 7 Spectrogram, waveform and beats for Example (15). Beats placed automatically by algorithm (<http://csppeech.ucd.ie/~fred/beatExtraction.php>).

A, and projects an up-coming, on-beat turn. B’s turn is indeed aligned in time and in action with A’s. The isochrony across the two turns at talk (Couper-Kuhlen 1993, Auer, Couper-Kuhlen & Müller 1999) is thus prosodically co-constructed, in a manner similar to syntactic co-constructions (e.g. Lerner 2004).

While Example (15) showed a click projecting an on-time and on-beat response, the next example of clicks as metronomes involves clicks produced in one turn being used as a device for incoming laughter.

(16) Salford A & R lateral click 1654 topless beaches

```

01 A      I'll just go on the tOpless `BEAches;
02 R      yOU -WISH-
03 A      -nO- ?I know whEre they `ARE,
04        | | | [|
05 R      [ <<laugh> * * * * * * * * * * * * * * ]
06 A      †And †Itty bars,
07 R      rIGHt;=Anyway.
08 A      <<laugh>- [----->]
09 R      [so where Else would <<laugh> you wanna `GO.>]
    
```

Anthony and Ray have been talking about places they would like to go to. Anthony has mentioned that he would like to go to Australia. Ray, in line 2, dismisses Anthony’s plans, and Anthony in line 3 rejects Ray’s rejection. Anthony’s TCU is then extended by a series of four lateral clicks, which are isochronously spaced. These clicks have incrementally falling secondary resonances, probably achieved by progressive lip rounding. This change in resonance may be part of how a series of four is projected. Ray displays his understanding of Anthony’s response as something laughable. Ray’s laughter is pulsed, and the first pulse comes in on beat with Anthony’s last click. Since the intervals between Anthony’s four clicks are equidistant, the clicks set up a fast beat, which Ray orients to with his turn

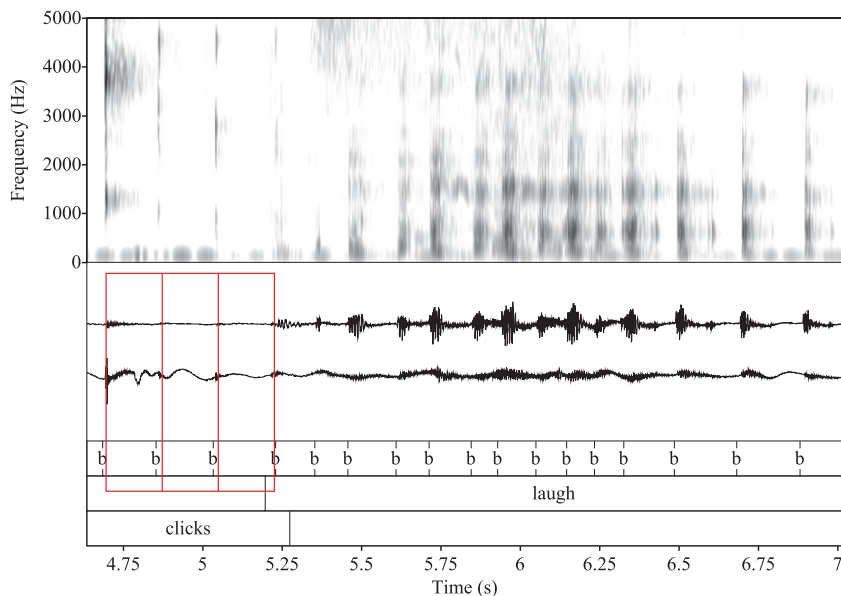


Figure 8 Example (16): Multiple clicks followed by laughter from another speaker. Beats are placed on each of the clicks, and the final click coincides with the onset of the laughter pulses. The red grid is aligned with the clicks, and the lines are spaced without reference to the audio. The laughter pulses are initially timed with shorter intervals than the clicks, but start on beat with the final click.

that consists of laughter: this laughter constitutes an appreciation of Anthony's rebuttal of Ray's rejection in line 2. Figure 8 shows Anthony's clicks, again with beats marked automatically.

Anthony's laughter pulses are spaced more closely together than Anthony's clicks, but nonetheless, he displays an orientation to the beats set up by Anthony's clicks by the way he times his incoming laughter, which displays an appreciation of Anthony's prior turn. These two examples (and others like them) show that clicks can simultaneously handle matters of interpersonal timing in talk and issues of stance-taking. Because they are punctual, clicks provide participants with clear signals about rhythm and timing in talk. It is known (e.g. Auer et al. 1999) that the temporal location of incoming talk affects the interpretation of turns at talk. Clicks can serve as a resource – a metronome – for participants (both speakers and hearers) to either solicit or to provide a well-timed and affiliating response.

9 Conclusions

Overall, our study has shown that clicks occur in (at least) three distinct systems with distinct functions but partially overlapping phonetic exponents. The details of these systems are not complete, but in outline they seem to be:

1. **MARKING INCIPIENT SPEAKERSHIP.** Clicks in turn-initial position can serve to project speech, and therefore can mark the transition from listener to incipient speaker. This system includes percussives which arise from the separation of the articulators, often simultaneously with inhalation. The sounds in this system are united by the features of

audible separation of articulators, and ingressive airflow, which may be either pulmonic (percussive) or velaric (click).

2. ASPECTS OF SEQUENCE MANAGEMENT. Clicks and percussives both commonly occur as elements of word-search, in conjunction with other, more linguistic features: particles such as *uh*, *uhm*, *mmm*, uttered on a mid-level pitch, perhaps tagged on to (full forms of) words; in-breaths and swallowing; gaps in talk at points where longer syntactic units are projected. Clicks also index new sequences of talk, at different syntactic locations from word searches. These uses of clicks, which appear to be more linguistically structured and particular, may have evolved from those sounds which mark incipient speakership, which appear more organic.
3. DISPLAYING A STANCE. Clicks, but apparently not percussives, are part of a family of verbal and non-verbal practices for displaying a stance turn-initially. Such clicks tend to be loud and deliberate, giving the impression that they are designed for a recipient to hear (unlike at least some sounds marking incipient speakership). They are commonly, but not necessarily, followed by response tokens whose prosodic design carries a great deal of communicative weight which we have not examined in this paper. Stance-taking clicks may also be accompanied by other bodily activities, such as head-turns, and probably facial expressions too. They are embedded in sequences of talk which can project the stance displayed in the turn with the click, and/or elaborated on after the click is produced, e.g. through overt linguistic material.

While we have not considered other languages, system 1, ‘incipient speakership’, may be subject to individual variation due to differences in e.g. amount and texture of saliva, and perhaps personal habits such as ‘neutral’ mouth postures (Scobbie et al. 2011). We might therefore expect to see similar patterns in other languages. On the other hand, word searches, new sequence indexing, and displays of stance involve more overtly linguistic elements, so we might expect to find significant differences between languages in terms of form and function. This remains a question for further research (though Pohjola 2011 has very similar findings for Finnish conversation).

Clicks, as the lay interpretation assumes, are involved in displays of stance. But the form of clicks in different kinds of stance-taking does not seem to be very variable. Rather, the variable elements are the location of clicks in a sequence of talk, and the accompanying response tokens. In other words, it is hard to demonstrate that clicks by themselves display a stance. It is much easier to show that turns displaying a stance occur in sequential locations where displays of stance are relevant, and that the kind of stance displayed is projected through the talk itself. Indeed, we have seen examples of both broadly positive and negative stance-taking performed with the click [!]. This indicates that in order to understand the linguistic components of stance-taking better, it is important to look at a longer sequence; and to (literally) look at non-verbal behaviour as well.

In summary, clicks and percussives in English have a complex but regular distribution. It is true that they are not phonemic, but there are good arguments for treating them as part of a speaker’s communicative competence. They have a regular distribution in pre-turn position (or mid-turn position in the case of word searches); and they are elements of bigger meaning-bearing prosodic constructions. It seems very likely that there are individual variations in the precise forms of clicks, and the frequency with which they are used. For more conclusive statements about their linguistic value, or their specific functions, cross-linguistic studies are needed.

The analysis in this paper is deficient in that it depends heavily on auditory and acoustic analysis. This is necessarily so, as the competing demands between natural data – where participants have a real social, interactional and emotional investment in what they are doing through their talk – and laboratory data – where the social situation is highly controlled – are difficult to reconcile. Articulatory studies would provide much more detailed information about the precise articulation of clicks and percussives, and may help us to explore in better

detail the acoustics of such sounds. Access to video data indicates that a fuller study could also usefully consider the body more generally: breathing and swallowing are two basic activities that are intimately bound up with clicks and percussives in the data considered here. Securing such data, while retaining natural contexts, is however extremely challenging. A useful next step would be to conduct laboratory studies so that the relation between articulation and acoustics is clearer, in order to make acoustic analysis of natural data, alongside a careful sequential analysis, more robust.

Acknowledgements

Thanks to Daniel Clubb, Marianna Kaimaki, and Emma Piggott for help with data collection; to Bill Haddican and Hazel Richards for access to the ESRC Northern Englishes data; to Emanuela Buizza for help with the illustrations; to anonymous reviewers of JIPA whose comments helped to improve the paper; to Erica Gold, Auli Hakulinen, Sara Howard, Tamar Keren-Portnoy, John Local, Anna Pohjola, Jim Scobbie, Marja-Leena Sorjonen, Jane Stuart-Smith, Jürgen Trouvain, Marilyn Vihman, Melissa Wright and many other colleagues for useful discussions along the way. The University of York Research Priming Fund supported some of the data collection and analysis.

Appendix: Data sources

CallHome	Recordings made in the USA, approximately 30 minutes in duration. The caller received a free 30-minute call in return for allowing the call to be recorded. Most calls are to friends and family outside the USA.
SW/CK/Sym	A collection of displays of sympathy (Wilkinson & Kitzinger 2006).
ESF/O	A collection of naturally-occurring offers taken from a wide range of Conversation Analytic sources, including other corpora named here.
Holt	Telephone recordings made in the 1980s in a British family home as part of a conversation analytic corpus.
lists	(preceded by initials, ell lists, klm lists). A collection of approximately 350 naturally-occurring lists collected by students from radio phone-in shows, mostly drawing on local radio programmes in the United Kingdom.
nrb	Recordings made in a recording studio between students. The participants were asked to catch up with one another's news, but were given few other instructions.
RCE25	Part of a corpus of video data of conversations at the University of York. The data used in this paper are from two friends catching up with each other.
Salford	Sociolinguistic interviews as part of a project on Northern Englishes (ESRC REF).
SBL	Recordings made in California, focusing on a circle of female friends. Date uncertain; 1960s.
vegtalk	A BBC radio phone-in show from the early 2000s. Each show has two presenters, and usually an expert in the studio.
Virginia	An early video recording in black and white of a family dinner in the United States in the late 1950s/early 1960s (date uncertain).

References

- Abercrombie, David. 1967. *Elements of general phonetics*. Edinburgh: Edinburgh University Press.
- Auer, Peter, Elizabeth Couper-Kuhlen & Frank Müller. 1999. *Language in time: The rhythm and tempo of spoken interaction*. Oxford: Oxford University Press.

- Ball, Martin. 1989. *Phonetics for speech pathology*. London: Taylor and Francis. [Reissued 1990, London: Whurr.]
- Button, Graham. 1987. Moving out of closings. In Graham Button & J. R. Lee (eds.), *Talk and social organisation*, 101–151. Clevedon: Multilingual Matters.
- Catford, J. C. 2001. *A practical introduction to phonetics*. Oxford: Oxford University Press.
- Clark, John & Colin Yallop. 1990. *An introduction to phonetics and phonology*. Oxford: Basil Blackwell.
- Couper-Kuhlen, Elizabeth. 1993 *English speech rhythm: Form and function in everyday interaction*. Amsterdam & Philadelphia, PA: John Benjamins.
- Crystal, David. 1987. *The Cambridge encyclopaedia of language*. Cambridge: Cambridge University Press.
- Cummins, Fred & Robert Port. 1998. Rhythmic constraints on stress timing in English. *Journal of Phonetics* 26(2), 145–171.
- Drew, Paul. 1998. Complaints about transgressions and misconduct. *Research on Language & Social Interaction* 31(3), 295–325.
- Ford, Cecilia E. & Sandra A. Thompson. 1996. Interactional units in conversation analysis: Syntactic, intonational and pragmatic resources for the management of turns. In Elinor Ochs, Emanuel A. Schegloff & Sandra A. Thompson (eds.), *Interaction and grammar* (Studies in Interactional Sociolinguistics 13), 134–184. Cambridge: Cambridge University Press.
- Gil, David. 2011. Para-linguistic usages of clicks. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Munich: Max Planck Digital Library.
- Gimson, A. C. 1970. *An introduction to the pronunciation of English*, 2nd edn. Bristol: Edward Arnold.
- Goodwin, Charles. 1981. *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press.
- Goodwin, Charles. 1986. Gestures as a resource for the organization of mutual orientation. *Semiotica* 62(1–2), 29–49.
- Gussenhoven, Carlos. 2004. *The phonology of tone and intonation*. Cambridge: Cambridge University Press.
- IPA [International Phonetic Association]. 1999. *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press.
- Jefferson, Gail. 1984. On the organization of laughter in talk about troubles. In J. Maxwell Atkinson & John Heritage (eds.), *Structures of social action: Studies in conversation analysis*, 346–369. Cambridge: Cambridge University Press.
- Kohler, Klaus J. 2007. “Speech–smile”, “speech–laugh”, “laughter” and their sequencing in dialogic interaction. *Interdisciplinary Workshop on the Phonetics of Laughter*, Saarbrücken, 21–26.
- Ladefoged, Peter. 1982. *A course in phonetics*, 2nd edn. New York: Harcourt Brace Jovanovich.
- Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford: Blackwell.
- Ladefoged, Peter & Anthony Traill. 1994. Clicks and their accompaniments. *Journal of Phonetics* 22, 33–64.
- Laver, John. 1994. *Principles of phonetics*. Cambridge: Cambridge University Press.
- Lerner, Gene H. 2004. Collaborative turn sequences. In Gene H. Lerner (ed.), *Conversation analysis: Studies from the first generation*, 225–256. Amsterdam: John Benjamins.
- Local, John & Gareth Walker. 2008. Stance and affect in conversation: On the interplay of sequential and phonetic resources. *Text & Talk – An Interdisciplinary Journal of Language, Discourse Communication Studies* 28(6), 723–747.
- Norrick, Neal R. 2012. Listening practices in English conversation: The responses responses elicit. *Journal of Pragmatics* 44(5), 566–576.
- Ogden, Richard. 2010. Prosodic constructions in making complaints. In Dagmar Barth-Weingarten, Elisabeth Reber & Margret Selting (eds.), *Prosody in interaction*, 81–104. Amsterdam: John Benjamins.
- Ogden, Richard. 2012a. The phonetics of talk in interaction – introduction to the Special Issue. In Richard Ogden (ed.), *The phonetics of talk-in-interaction*, special issue of *Language and Speech* 55(1), 3–11.
- Ogden, Richard. 2012b. Making sense of outliers. *Phonetica* 69(1–2), 48–67.
- Pike, Kenneth. 1943. *Phonetics: A critical analysis of phonetic theory and a technic for the practical description of sounds*. Ann Arbor, MI: University of Michigan Press.

- Pohjola, Anna. 2011. Maiskaukset keskustelussa [Clicks in conversation]. BA dissertation, Finnish Language, University of Helsinki.
- Potter, Jonathan & Alexa Hepburn. 2010. Putting aspiration into words: “Laugh particles”, managing descriptive trouble and modulating action. *Journal of Pragmatics* 42(6), 1543–1555.
- Reber, Elisabeth. 2012. *Affectivity in interaction: Sound objects in English*. Amsterdam: John Benjamins.
- Sacks, Harvey, Emanuel A. Schegloff & Gail Jefferson. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50, 696–735.
- Schegloff, Emanuel A. & Gene H. Lerner. 2009. Beginning to respond: Well-prefaced responses to *wh*-questions. *Research on Language and Social Interaction* 42(2), 91–115.
- Schegloff, Emanuel A. & Harvey Sacks. 1973. Opening up closings. *Semiotica* VIII(4), 289–327.
- Scobbie, Jim M., Sonja Schaeffler & Ineke Mennen. 2011. Audible aspects of speech preparation. *17th International Congress of the Phonetic Sciences (ICPhS XVII)*, Hong Kong, 1782–1785.
- Selting, Margret, Peter Auer, Dagmar Barth-Weingarten, Jörg Bergmann, Pia Bergmann, Karin Birkner, Elizabeth Couper-Kuhlen, Arnulf Deppermann, Peter Gilles, Susanne Günthner, Martin Hartung, Friederike Kern, Christine Mertzluff, Christian Meyer, Miriam Morek, Frank Oberzaucher, Jörg Peters, Uta Quasthoff, Wilfried Schütte, Anja Stukenbrock & Susanne Uhmann. 2009. Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 10, 353–402.
- Ward, Nigel. 2006. Non-lexical conversational sounds in American English. *Pragmatics and Cognition* 14(1), 113–184.
- Wells, Bill & Sarah MacFarlane. 1998. Prosody as an interactional resource: Turn-projection and overlap. In Marc Swerts & Julia Hirschberg (ed.), *Prosody and conversion*, special issue of *Language and Speech* 41(3–4), 265–294.
- Wilkinson, Sue & Celia Kitzinger. 2006. Surprise as an interactional achievement: Reaction tokens in conversation. *Social Psychology Quarterly* 69(2), 150–182.
- Wright, Melissa. 2005. *Studies of the phonetics–interaction interface: Clicks and interactional structures in English conversation*. Ph.D. thesis, University of York.
- Wright, Melissa. 2007. Clicks as markers of new sequences in English conversation. *16th International Congress of the Phonetic Sciences (ICPhS XVI)*, Saarbrücken, 1069–1072.
- Wright, Melissa. 2011a. On clicks in English talk-in-interaction. *Journal of the International Phonetic Association* 41(2), 207–229.
- Wright, Melissa. 2011b. The phonetics–interaction interface in the initiation of closings in everyday English telephone calls. *Journal of Pragmatics* 43(4), 1080–1099.