Escalating prosody: the vocal depiction of mobile action sequences

Beatrice Szczepek Reed, King’s College London

Abstract

This paper shows how horse-riding instructors vocally depict, and thereby co-design and mobilize the unfolding progression of horses’ and riders’ actions through gradually escalating prosody. Escalating prosody involves the stepwise raising of speakers’ overall pitch across a series of turn components, often accompanied by increases in overall loudness, and occasionally by changes in voice quality. Escalating prosody can accompany instructed activities from beginning to end, or only during certain phases of the activity. The prosodic delivery mirrors the building and subsequent sustaining of physical effort expected of the horse-rider pair. It can occur with lexical instructions to perform series of actions or with repeated directives to sustain the current activity. It can also occur with repeated praise as a successful performance unfolds, and with repeated corrections, which temporally frame moments of trouble. Prosody is shown to be a resource for co-designing the actions of others, specifically, their mobility, sustained-ness, and sequential progression. The data are horse-riding lessons recorded in the UK and Germany.

Keywords: interactional linguistics, intonation, instruction, horse-riding

1. Introduction

Prosody plays an important role in the management of talk-in-interaction, including turn-taking, action formation, sequence organization, recipient design, and the display of affect and stance. Under-researched areas are the contribution of prosody to the management of
activities as they unfold, and to the co-construction of others’ bodily actions. The present study shows that in one-to-one embodied skills instruction, the gradual building of other participants’ embodied activities is co-constructed not only lexically, through verbal directive-giving, but also prosodically, through a gradual build-up in vocal delivery. This is referred to here as ‘escalating prosody’ and includes increases in pitch register\(^1\), loudness, and changes in voice quality (for example, increasing harshness or increasingly tense creak). A first extract from a horse-riding lesson demonstrates escalating prosody.

(1) 20220525 2:9.34

1  I:  \textit{GOOD:D, (0.8) }+(0.9)+ +\textit{GOOD:D, (1.4)}
H/R:  \textit{>>canter-}>
H/R:  +\textit{turn-} +\textit{straight-}>

2  -> I:  \textit{PERfect;} (0.3) \textit{PERfect;} (0.3) \textit{<<f> PERfect-} (0.3)
H/R:  \textit{&cross half-way point&}

3  -> \textit{<<ff>} \textit{PERfect;} (0.3) \textit{PERfect-} (0.3)
H/R:  \textit{&cross half-way point&}

   \begin{center}
   \includegraphics[width=0.5\textwidth]{image1.png}
   \end{center}

4  I:  \textit{+that FEEL GOOD?}
H/R:  +\textit{straight-}>

Horse and rider have been practicing coming down the middle of the riding arena in a straight line. In the transcribed extract, they do so successfully. As the straight line unfolds, the instructor receipts their successful performance with a series of repeated positive assessment tokens, initially with two iterations of \textit{good} (line 1). As horse and rider approach and then cross the half-way point, which means that they have successfully accomplished about half of

\(^1\) The term ‘pitch register’ refers to the locally occasioned raising or lowering of a speaker’s overall pitch across turns or turn components, in contrast to ‘pitch range’, which describes a speaker’s overall pitch across a given interaction (Szczep Reed 2010).
the exercise, the instructor begins to assess their performance repeatedly with perfect as they gradually progress down a straight line. The item is first delivered with a pitch register of 200-226 Hz. Its first repetition rises to 232-258 Hz, the second to 268-307 Hz, and the third to 290-500 Hz, shown in Figure 1. In addition, Figure 2 shows increases in amplitude on the last two iterations (the first two iterations are accompanied by wind noise). Each iteration of perfect is delivered within the overall temporal span of two horse strides and with near-equal temporal distance (approx. 0.3s). The prosodic delivery combines with lexical repetitions to depict the horse’s and rider’s mobile progression towards, and ultimate achievement of, the instructed task.

Fig 1. Pitch trace, Extract (1), lines 2-3.
2. Sound patterns in instruction sequences

The study reported here contributes to the growing field of interaction analysis of embodied skills instruction (e.g., Ehmer & Brône 2021). Research that focuses specifically on the prosodic design of instruction talk is comparatively rare, but some previous studies have...
explored the affordances of different sound patterns for depicting instructional matters and for synchronizing, shaping, and enacting the bodies of others. For example, two studies by Keevallik (2014; 2021) show how dance instructors can accompany their demonstrations of dance moves with non-lexical vocalizations, thus depicting the rhythm and choreography of the dance. Albert and vom Lehn (2023) and Grahn, Lindholm and Huhtamäki (2023) show that not only instructors but also learners use non-lexical vocalizations, for example, to index trouble or to coordinate their movements. The phonetically underspecified nature of non-lexical vocalizations (Keevallik & Ogden 2020) provides opportunities for the creative depiction of purely embodied content, and for fitting vocalizations to actions without the constraints of lexis and grammar.

Like non-lexical vocalizations, prosody affords speakers a temporal and depictive flexibility that can be used to communicate embodied actions and align them with relevant temporalities. For example, Keevallik’s (2020) study of a Pilates session shows that instructors can synchronize the group’s embodied actions through repeated directives whose final iteration is produced with slower, louder, and higher prosody. This prosodic pattern signals the end of the exercise and brings the last group member on board. In another study of a Pilates lesson, Hofstetter and Keevallik (2023) investigate the prosodic adjustment and synchronization of others’ bodies. The authors identify an ‘UP/DOWN contour pair’ (p. 55), which accompanies embodied demonstrations by the instructor; for example, ‘and to one side? And to the other side.’ (p. 57). With the ‘UP/DOWN’ pitch pattern, the instructor indexes the chronological structure and rhythm of the exercise and again brings on board group members who are not yet aligned. Ehmer’s (2023) study of Tango lessons also finds a rising – falling pitch pattern on bipartite instructions, such as ‘forward – turn around’, which, he argues, contributes to the recognizability of instructions. While these studies show how prosodic patterns can display the structure of an instructed activity, Simone’s and Galatolo’s
(2021) investigation of rock-climbing lessons for visually impaired climbers finds that prosody can iconically represent the directives it accompanies, such as sound lengthening on the word ‘stretched’ (p. 402). Finally, prosody can index urgency in the face of delayed learner compliance. Mondada (2017) shows that fast repetitions of the same directive during race-driving instruction contribute to the co-construction of urgency and put ‘temporal pressure’ (p. 87) on the learner until the instructed action is accomplished (see also Mondada 2018a). Similarly, Okada (2023) shows that boxing coaches can target boxers’ delayed compliance with repeated imperatives that become faster, louder, longer in duration, or higher in pitch.

In the above work, prosody is shown to be a resource with which instructors co-design diverse temporalities of instructed actions (rhythm, synchrony, chronological structure, duration, urgency). The study presented here adds to this the prosodic co-construction of extended, to-be-sustained embodied activities, whose temporal unfolding and increasing bodily effort is depicted through gradually building prosody. With regard to the second aspect, the vocal depiction of increasing bodily effort, two studies are relevant that show prosody to be a resource for the enactment of one’s own and others’ bodies. Brandenberger and Hottiger (2018) analyze the talk of a person who is separating a set of disks from a magnet, an activity that is part of a visit to a science museum. Prosodic emphasis on a stressed syllable co-occurs with the moment preceding the separation, ‘RICHT(.)ig’ (German ‘right’, p. 65). An intra-word micropause appears to depict the moment of release. While this observation concerns the prosodic depiction of a speaker’s own bodily actions, Keevallik et al. (2023) have shown that in a range of interaction types, including dance and Pilates instruction, participants vocalize the embodied experiences of others through non-lexical vocalizations and their phonetic and prosodic design. Experiences that are being enacted in this way include pain, strain, and tasting.
This body of evidence shows that the interactional affordances of linguistic sound patterns go beyond the management of turns, sequences, and affect displays. They also include synchronizing, depicting, and enacting the bodies of others. The instruction of embodied skills appears to be a setting where these are key concerns for participants, as instructors share with learners the responsibility for the successful accomplishment of embodied outcomes. The analysis below shows that during embodied instruction, sound patterns can manage the progression of others’ mobile activities, especially those that involve moving towards, reaching, and maintaining a peak in their performance.

3. Intensified prosodic delivery

Escalating prosody involves a gradual move towards more marked, intensified prosodic delivery. ‘Intensifying emphasis’ has been described by Ogden (2012) for individual words, such as quantifiers, extreme case formulations, and assessments, which are treated by recipients as delivering out of the ordinary information. More gradual intensification over longer stretches of talk, which is the focus of this study, has been found in contexts where speakers depict the rising drama and suspense of an unfolding event. Studies investigating the prosody of sports commentary have described increases in pitch, tempo, and loudness, amongst other features. In an investigation of British English horse-racing TV commentary, Trouvain and Barry (2000) compare findings from three studies on the prosody of excitement (Banse & Scherer 1996; van Beezooyen 1984; Murray & Arnott 1993) to British racing commentators’ prosody. Horse-racing commentary is shown to involve a gradual build-up of pitch and loudness from start to finish of the race and an immediate decrease in both immediately afterwards. The strong increases in pitch and loudness at the climax of the race show similarities with those that have been claimed for displays of anger, joy, and surprise.
Several studies have also examined the prosodic features of football commentary. Kern (2010) finds that German radio commentators build suspense through a gradual rise in the overall pitch of intonation units (see also Müller 2007). Additional features are high rising pitch on the first intonation unit, a narrow pitch span within units, level pitch on the final intonation unit, an absence of pauses with frequent latching of intonation units, and an absence of lengthening. Commentary on goals is delivered with higher-than-previous pitch, pauses within intonation units, and vowel lengthening. Trouvain (2015), examining the historical development of German football commentary, also describes ‘a continuous or stepwise elevation of the pitch registers by between one and two octaves’ during pre-goal suspense talk; as well as high-pitched climaxes in response to goals.

Sports commentary can involve the dramatic build-up towards a projected endpoint. This build-up can be depicted prosodically. The analysis below will show similarities between sports commentary and the prosodic co-design of unfolding activities in horse-riding lessons, especially with regard to the gradual rise of overall pitch and loudness. While sports commentators’ prosody depicts in-progress developments for an overhearing audience, the instruction data presented below show sound patterns whose role is to shape, direct, and mobilize the bodies of others.

4. Data

The data for this study come from a corpus of 18 one-to-one horse-riding lessons. 16 are in English and were recorded in the UK between March 2022 and March 2023 (9 hours 8 minutes); 2 are in German and were recorded in 2022 (1 hour 29 minutes). For the UK corpus, the author approached 21 riding coaches with a request to film their private lessons. 18 were contacted through the webpage of the British Horse Society, and three agreed to take
part in the study. Three additional coaches were approached personally and are known to the author; all three also agreed to be recorded. The coaches and their clients gave written consent to be audio and video recorded, and for their anonymized data to be published for the purpose of research dissemination. The study received ethical approval from King’s College London in March 2022.

The two German recordings are available on the ‘Database for Spoken German’ of the Institute for the German Language (Datenbank für gesprochenes Deutsch, Leibniz-Institut für Deutsche Sprache), specifically the FOLK corpus (Schmidt 2014). The two recordings can be found under FOLK_REIT_01_A01 and FOLK_REIT_01_A02. These recordings were included because in them, instructors and riders communicate via their mobile phones: the instructor holds her mobile phone on her lap; it is switched to speaker phone. The rider listens and speaks via earphones. Since the object of this study is heightened prosodic delivery, the author was keen to observe whether the absence of the need to project the voice across a physical distance impacts on the frequency or format of escalated prosody. Due to the small sample size of technology-assisted riding lessons, and due to the fact that no other variables were held constant in the collection of the overall data, any observations on this issue will be anecdotal at this stage but could inform future research.

The collection contains 25 clear cases of escalating prosody. The verbal transcription follows GAT2 (Selting et al. 2009/2011, see Appendix), except that intonation units have not been allocated individual lines in the transcripts. In order to maintain a manageable transcript length, intonation units are divided by unit-final punctuation marks only, but transcript lines can contain more than one unit. The embodied transcription follows Mondada (2018a, 2019). Frequency analyses and wave forms were produced in PRAAT 6.2.07 (Boersma & Weenink 2022); figures showing pitch or amplitude are labelled ‘Figure’. Still pictures from video recordings are labelled ‘Image’. All transcripts from the UK collection include at least one -
anonymized - image to give a sense of the physical setting. The transcripts cover only those aspects of the interaction that are relevant to the analysis. The recordings show many additional details, and transcripts are understood to be ‘unavoidably incomplete’ (Ten Have 1997) records of recordings, twice removed from the original event (Ashmore & Reed 2000).

5. An interactionalist approach to horse-riding lessons

The study aligns with the conceptual and methodological approaches of Conversation Analysis (Schegloff 2007) and Interactional Linguistics (Couper-Kuhlen & Selting 2018). This means that it treats language phenomena as resources for social interaction. A key concern is the emergent, sequentially ordered development of interactional events. The relevance of events to the analysis is not assumed but must be evidenced by participants’ treatment of them as relevant, typically through their verbal and/or bodily responses. Horse-riding lessons present a challenge in this regard, both in terms of the human and nonhuman participants. Firstly, most embodied actions by riders cannot – and should not – be visible. The desirability of the subtleness of the riding aids is a ‘members’ problem’ (Hutchby & Wooffitt 2008, p. 50), since riding instructors, too, can often not see what riders are doing with their bodies. According to one participant in this study, they often infer riders’ actions from the way horses move, rather than from watching riders’ bodies. The research presented here is careful to focus only on instances where clearly visible actions are being made relevant by participants. Secondly, the analysis mentions horses’ responses only when these are clearly done in relation to riders’ or instructors’ actions.

Another feature of horse-riding lessons is that most instructor talk is designed only for the human recipients, despite the joint accomplishment of ridden actions by horses and riders (for instructor talk designed for both horses and riders see Szczepek Reed (2023a)). This disparity
is addressed here by referring to actions by the ‘horse-rider pair’, and to horse- or rider-only actions when these are clearly performed individually.

Previous interaction-oriented work on horse-riding lessons has investigated how instructors teach the subtle communication between horse and rider (Norris 2012; Lundgren 2020; Lundesjö Kvart & Melander Bowden 2022) and on instructors’ communicative practices (Lundgren 2017; Lundesjö Kvart 2020). Three previous studies by the author have explored how riding instructors use prosody and nonlexical vocalizations to design their talk for horses and riders at the same time (Szczepek Reed 2023a); how they teach ‘sustained embodied activities’, which are a typical feature of horse-riding instruction (Szczepek Reed 2023b, see next section); and how they deploy silent observation in a mostly talk-based instruction context (Szczepek Reed in press).

6. Escalating prosody: depicting the mobile activities of others

The following sections show how escalating prosody accompanies instructed embodied activities as they unfold. Instructor talk that emerges concurrently with learner performances has been the topic of a range of studies, including work on driving instruction (De Stefani & Gazin 2014; Deppermann 2018a, b; Mondada 2017, 2018a), dance instruction (Ehmer 2021; Reed forthcoming), boxing coaching (Okada 2018; 2023), Pilates lessons (Keevallik 2020; Hofstetter and Keevallik 2023), rock-climbing instruction (Simone & Galatolo 2020, 2021), and horse-riding lessons (Lundesjö Kvart & Melander Bowden 2022). In such settings, a distinction is relevant between instructions that are to be complied with now, in the slot following an instructor’s directive; and those that apply more generally and/ or are to be done at an unspecified time in the future. Vine (2004, pp. 32-33) calls these ‘NOW’ vs. ‘LATER directives’, while Szczepek Reed, Reed and Haddon (2013) distinguish between ‘local’ and
‘non-local action directives’. This distinction will be relevant below, as prosodic escalation can be found only with local NOW directives. The following three sub-sections show escalating prosody with the instruction of to-be-sustained activities, with repeated praise, and with repeated corrections.

6.1 Instructing to-be-sustained activities

Riding instruction often involves longer activities that have to be sustained over an extended period. In Szczepak Reed (2023b) I have described ‘sustained embodied activities’, that is, activities that require a continuing physical effort as well as series of embodied actions that are performed simultaneously. Examples are asking the horse to walk sideways for several strides or working on a specific gait (walk, trot, canter). That analysis shows how riding instructors orient to a range of temporalities when teaching these complex activities. Most relevant to the study presented below is their verbal co-construction of a continuing timeline, since the very act of sustaining such activities is an object of instruction. Practices for doing so include directives such as go on, keep going, and and again, which do not target a specific embodied action but instead request the continuation of whatever action is currently being performed. Practices for co-constructing sustained activities also include directives to perform certain actions repeatedly and from now on, for an unspecified number of times, which opens up an extended timeline during which these directives remain relevant; and the repeated use of rise-to-mid intonation on individual directives, which design instructed actions as components of an ongoing activity. As that study shows, in an instructional context it is not only the quality of such activities that is under scrutiny but also their very sustained-ness: keeping the activity going can in itself be treated as an achievement.
When riding instructors teach to-be-sustained activities, their prosody not only designs the activity as one that continues, but it can also reflect the gradual build-up of effort that horses and riders are being asked to invest. This can be observed clearly in 16 instances in the corpus. The prosodic features include a stepwise rise in pitch register, often accompanied by increasing loudness, and occasionally by harsh or tense voice quality on later items. Extract (2) shows a sequence from a beginner riding lesson in which a young rider is being taught how to ask a horse to canter, which is the three-beat gait up from the two-beat gait of trot.² The instructor’s escalating prosody accompanies the horse’s and rider’s actions from the onset of the preparatory phase through to the end of the instructed activity. The symbol ǁ represents lateral clicks (see Ogden 2020).

(2) 20220425 Lesson 1 1:13.58

1 –> I:  *that’s #IT- hands* DOWN, keep his HEAD straight=*
  H/R:  >>trot>+
  H:  *turns head to right*
  Image #Image1

2 –> I:  =keep that left REIN O:N?%* (0.3)
  H/R:  %through corner->%

3 –> I:  right hold your saddle <<f> NOW?:% (0.9)%
  H/R:  ---------------------------------------------------%

4 –> I:  %<ch+f> SIT=>% & (0.4) <<h+f+harsh> KICK, % & (0.5)+ #:CANTer.>
  H/R:  ---------------------------------------------------+ +canter->+
  H/R:  %approach C--% %pass C---------------------% %through corner->%
  R:  &sits->& &leg

² The extract contains repeated directives to *kick*, which as a term appears very forceful. Riding involves the use of the human’s legs against the body of the horse; however, this recording is the only one in the corpus where the term ‘kick’ is being used. In all other lessons, the instructors – including the one shown here - use phrases such as ‘use your leg’, ‘put the leg on’, or ‘nudge with the leg’. All other lessons are with adult riders. It is possible that in using a more forceful term here the instructor is orienting to this rider’s younger age and lesser physical strength, and to her competence level as a beginner.
5  ->  I:  <<h+f> and ↑GO. #(0.6) Left hand DOWN and KICK.> & (0.8) #
H/R:  --------------------------------------------------%
R:  &leg

6  ->  I:  [&that’s IT- <<h+f> well &↓DONE.
R:  &leg &leg
H:  [stride [stride

7  ->  I:  keep ↑GOing. leg a↓[GAIN. (0.6) and a↓[GAIN. [↑ ] [↑ ]
R:  &leg &leg
H:  [stride [stride [stride

8  ->  I:  <<f+hh> and a↓[↑GAIN. # keep ↑GOing. [↑ ] [↑ ] and a↓[↑GAIN. [↑ ] [↑ ]
R:  &leg &leg
H:  [stride [stride [stride

9  ->  I:  and a↓[↑GAIN. [↑ ] [↑ ] [↑ ] <<extra h> THAT’S &↑↑IT?> (0.5)
10 R:  [.hhh &he
H:  [stride &leg
R:  &leg

11 I:  get rEAdy for your &TROT, (0.5) pull your REIN, (0.9)
R:  &leg

12 I:  then RI++:SE. &
H/R:  ---------trot-->
R:  -----------&rising trot->>

14
In beginner lessons, cantering is typically done as the horse moves into a corner, as the bending movement makes it easier for the horse to pick up the canter, which in turn makes it easier for the rider to ask for it. Earlier transitions to canter in this lesson have been done in the second corner of one of the short sides of the arena. The instructor’s prosody begins to build as she prepares the rider for the upcoming transition to canter, and as horse and rider are trotting down a long side. Trotting is the two-beat gait between walk and canter, and because of its increased speed and momentum it typically precedes the faster, three-beat canter in beginner lessons. The phrases that’s it – hands down – keep his head straight – keep that left rein on are produced with a gradually ascending overall pitch register, as each turn component steps up to a higher pitch. This is most clearly observable on the accented syllables it, down, head, rein (Fig. 3). The first phrase, that’s it, starts at 190 Hz, while the final phrase rein on starts at 304 Hz and ends in a long, rising pitch contour reaching 405 Hz. Continuing this trajectory, the next directive right hold your saddle now shows a long rising final contour on now, rising to 450 Hz (Fig. 4). On now the instructor’s voice becomes noticeably louder (Fig. 5).

The instruction components are, first, an assessment (that’s it), and then a series of directives that need to be complied with now, in the moment of the emerging activity. hands down is likely to be a response to the rider’s hand position, although their hands cannot be seen in the video. The two directives keep his head straight – keep that left rein on target the fact that the horse has briefly turned his head to the right (line 1, Image 1). Contact with the outside rein (here, the left one) will also be needed for the transition to canter. The next directive, hold your saddle now, is given as horse and rider move through the first corner on the short side of the arena and are approaching the second corner, where they will canter.

The following directives sit kick canter represent the culmination of the rider’s preparatory actions which will end in the start of the instructed activity, the transition to canter. On sit, the
instructor steps up further, to approx. 460 Hz, kick remains slightly below, canter steps slightly above (Fig. 6). The three items have increased loudness, which cannot be represented visually due to the instructor’s distance from the microphone. In addition, sit and kick are produced with harsh voice quality. The pitch patterns resemble those of a closed list (Selting 2007): sit is delivered with level pitch, kick with a rise-to-mid, and canter with a fall-to-low.

Like Keevallik’s (2020) series of directives that end in slower, louder, and higher prosody, and like the UP/DOWN contour identified by Hofstetter and Keevallik (2023), this pitch pattern co-constructs the structure and progression of its recipients’ embodied actions.

The directive to sit is given as horse and rider approach the C marker, that is, the mid-point of the short side. ‘Sit’ refers to sitting trot: while the rider has been doing rising trot up to now, which means they have stood up from the saddle on every other trot stride, they are now being asked to remain seated, in preparation for the canter. The rider sits as they pass the C marker, and the instructor gives the directive to ‘kick’, which the rider complies with just as they enter into the second corner (Image 2). The horse picks up the canter straight away.

The instructor now sends the horse and rider forward with and go, and with an upward movement of her right arm (line 5, Image 3). The high pitch register and loudness are maintained during lines 5-7, with the exception of that’s it (line 6), which returns briefly to a less marked delivery. Following two positive assessments (that’s it well done) in response to the successful transition to canter, the instructor now begins a sequence of repeated directives that target sustaining the canter by using the leg: keep going leg again and again and again keep going and again and again (lines 7-9). From line 8 onwards, there is further prosodic escalation. Figure 7 shows the transition from line 7 to line 8. While and again at line 7 ranges from 335 – 587 Hz (indicated in the lefthand margin of the figure), and again at line 8 steps up to 457 – 671 Hz (indicated in the righthand margin). Again, the increase in amplitude cannot be shown due to the instructor’s distance from the recording device. Line 10 in the
transcript shows the rider exhaling loudly and producing a single laughter particle in response to the instructor’s escalating repetitions. In addition to the escalation of pitch and loudness, the accented syllables of all phrases from *that’s it* (line 6) to *and again* (line 9) are rhythmically synchronized with the horse’s canter strides, as the accented syllables align with the horse’s front right leg hitting the ground.

At line 9, there is a final escalation on the positive assessment *that’s it*. Figure 8 shows that the preceding two phrases *and again and again* maintain the pitch register established at line 8, ranging between 349 and 668 Hz (indicated in the lefthand margin of the figure). *that’s it* steps up to a starting pitch of 432 Hz and rises to 807 Hz (indicated in the righthand margin). These pitch values are very high indeed. Many claims have been made regarding men’s and women’s speaking ranges, and these vary greatly according to interactional contexts (Szczepk Reed 2010). However, a large-scale study on general population data in Germany found female speakers’ mean pitch for ‘shouting voice’ to be 246.2 Hz (Berg et al. 2017). In comparison, the speaker in Extract (2) already raises her pitch above this value at the very beginning of the prosodic escalation. Her final pitch, 807 Hz, is somewhere between a G₅ and a G#₅, that is, an octave plus a fifth above middle C; a high note in any soprano aria.

The assessment *that’s it* opens up the closing of the canter phase. In her next directive, the instructor asks the rider to *get ready for your trot*, that is, to prepare for the end of the instructed activity (line 11). Figure 8 shows the steep step down from the previous 807 Hz to 232 Hz on the first word *get* (indicated in the righthand margin).
Fig 3. Pitch trace, Extract (2), lines 1-2.

Fig 4. Pitch trace, Extract (2), line 3.
Fig 5. Wave form, Extract (2), line 3.

Fig 6. Pitch trace, Extract (2), line 4.
Fig 7. Pitch trace, Extract (2), lines 7-8.

Fig 8. Pitch trace, Extract (2), lines 8-11.
The extract shows how an instructed mobile activity is accompanied from beginning to end by gradually escalating prosody. The prosodic escalation progresses from the initial preparatory phase to a sustaining of the activity and through to its closure. Both verbally and prosodically, the activity is being framed as requiring bodily effort to bring it about in the first place, and then to maintain it. The increasing physical effort that is audible in the instructor’s raised pitch and amplitude mirrors the physical effort she requests of the rider: it mobilizes the rider to exert a similar degree of energy in asking the horse to canter and to keep cantering. During the sustaining phase, the instructor’s talk is additionally marked by lexical repetition and rhythmic synchronization.

Extract (2), above, shows that escalating prosody can accompany the preparation of an instructed activity as well as the instructed activity itself. In contrast, in the following extract, the instructor gives prosodically escalating directives at a point at which horse and rider have already started to perform the instructed activity. The rider’s task is to keep the horse cantering on a circle, which the horse has found hard to sustain. Prior to the transcribed section, the horse has repeatedly fallen from canter back into trot, maintaining canter only for a few strides at a time. Trot can be an easier gait for horses, as – in its natural form – it requires less physical effort. In the extract, the instructor asks the rider to use her inside leg in order to sustain the canter while the horse is on the circle. The phrase ‘on this rein’ (line 1) can be paraphrased as ‘when riding in this direction’.

(3) 20230327 1:23.02

1    I:    so you’re gonna Use your ↑LEG a bit more on this rein as wEll.
H/R:  >>canter at slow pace on a circle to the right←

2    → I:    #<<f+h> LEG?>  +||  ||  <<f+h+harsh> LEG?>  (0.5) <<f+h> aGAIN,
H:  +swishes tail
Image #Image1
The instructor gives the local NOW directive leg with high pitch (Figure 9), increased loudness, and high rising intonation. The directive is most probably targeted at the slow pace of the canter, which risks the horse falling back into trot. The rider’s response cannot be seen, but the horse swishes his tail which suggests she has touched him with her leg. The instructor clicks four times in an attempt to send the horse forward (Szczepek Reed 2023a), and then repeats her directive with increased loudness and pitch, and with harsh voice quality. She continues to give directives with high pitch and loudness, mobilizing the rider to sustain their physical effort to increase the horse’s tempo (again, again, (push) him on, leg, keep going). As the horse comes out of the corner and faces the long straight line of the riding arena, he makes an energetic canter stride forward and canters faster, now clearly committed to
cantering. The instructor’s voice drops immediately to her default pitch and loudness as she comments that the horse will find canter easier on a straight line (line 4, Image 2).

Like Extracts (1) and (2), this example shows how prosodic escalation depicts and co-designs the progression of a mobile activity. Like Extract (2), it also shows the mobilizing of physical effort to sustain an activity, which involves escalating prosody in combination with lexical repetition. The instructor’s vocal delivery shows no rhythmic synchronization with the horse’s movements, which indicates that their orientation is to the rider rather than to the horse-rider pair as a ‘mobile formation’ (McIlvenny, Broth & Haddington 2014).

![Fig 9. Pitch trace, Extract (3), lines 1-2.](image)

In the following extract from a German recording, a learnable has been practiced several times before the instructor uses escalating prosody with the final iteration. The task is to transition from walk, the slow four-beat gait, to canter, the faster three-beat gait, and back to
walk, without falling into trot, i.e. the two-beat gait between walk and canter. The exercise is to be done on a small circle. Cantering on a small circle can be hard work for a horse, and by extension, keeping the horse cantering can be hard work for the rider. By the time the extract begins, the exercise has already been performed 10 times. During several previous iterations, the horse has fallen back into trot after only a few strides of canter. Before the iteration preceding the transcribed extract, the instructor tells the rider to repeat the task one more time and then change direction. However, the horse falls back into trot after only two canter strides, upon which the rider announces that she will do ‘one more’ (data not shown). In response, the instructor coaches her through another iteration of the exercise, this time with escalating prosody.

(4) FOLK_E_00492 27.20

1 I: das is OKAY; das_s GUT; *(0.9)* || (0.5) JA: || (0.4)
   that’s okay that’s good // yes //
   H/R: walk small circle------* trot-* *canter small circle->

2 -> I: <<creak> KOMM- (0.3) KOMM- (0.3) KOMM- (0.3) <<tense> ; KOMM
   <<creak> come on come on come on <<tense> ; come on

3 -> I: (0.3) ↑↑KOMM- (0.3) ↑↑↑KOMM- (0.2) ↑↑↑und JA- (0.2) ↑↑↑JA->
   ↑↑↑come on ↑↑↑↑come on ↑↑↑and yes ↑↑↑↑yes

4 -> I: (0.4) und ↑↑↑↑JA:::;==und SCHRITT. (0.9) ↑↑↑↑GU::;T; (1.0)
   and ↑↑↑↑yes= =and walk ↑↑↑↑good

5 I: und HAND**wechsel.
   And change of rein
   H/R: -------**walk->>

The instructor supports the rider’s efforts to keep the horse going with repeated directives of *komm* (‘come on’) and with further encouragement *(und) ja* (‘(and) yes’). The recording quality does not allow for a reliable frequency analysis of the whole extract, but the directives can be heard to rise in pitch in a stepwise fashion. They also take on an increasingly tense voice quality in addition to the creak that is present across all items. The first three iterations
of *komm* are delivered at around the 300 Hz mark and with creak. The fourth *komm* steps up to a peak at 381 Hz and is produced with tense voice quality, which increases across the following iterations. The fifth *komm* peaks at 415 Hz, the sixth near 440 Hz. The following *und ja* and *ja* have clipped vowels and maintain the high pitch as well as tense and creaky phonation. The final *and ja* has a lengthened rise-fall pattern, peaking at approximately 500 Hz. The later item *gut* (‘good’) repeats this pitch contour. Interestingly, the intervening new directive to come back to walk (*und Schritt*, line 4) is not part of the prosodic escalation. The overall prosodic design is one of increasing pitch and vocal tension, culminating in a final fall. The repeated prompts construct the activity as one that requires strong and increasing effort as the horse canters for longer than on previous occasions. Unlike the previous extracts, the example does not show increases in loudness. This may be the case because the instructor communicates with the rider through her mobile phone. Like Extract (3), the extract also shows no rhythmic alignment with the horse’s canter strides.

When instructed actions are brief and near-instantaneous, the performance of the instructed action can occur before the final escalation and without verbal commentary by the instructor. Prosodically, these cases are always delivered with list-like intonation (Selting 2007), where non-final items end variably in rise-to-mid, level, or fall-to-mid pitch, and the final item ends in low falling pitch. In the following extract, a beginner rider is being asked to bring about a transition to canter.

(5) 20220724 Lesson 2 1:18.24

1  I:  so gEt your TROT eSTABlished and then; (.)
   H/R:  >>walk->+

2  I:  INside leg O:N;+ +OUTside lEg BACK, #(1.5)
   H/R:  -------------------+ +trot->+
   Image  #Image1
3 -> I: EYES ↑UP- #(1.9) #and then sit <<f>↑↑UP↑↑> (0.7)
H/R: -------------------------------++canter-->
Images #Image2#Image3

4 -> I: #<<ff> ↑↑↑GOOD. well ↑↑DONE.> #(0.4) and then try and sit
Image #Image4 #Image5

5 STILLer in your bOdy if you CAN,

The prosodic escalation occurs over a three-part structure: eyes ↑up – and then sit ↑↑up – ↑↑↑↓good, during which the individual items step higher until they end in a final fall-to-low.

The second and third item also increase in loudness; however, due to the strong wind noise on this recording, amplitude cannot be visualized.

At the beginning of the transcribed section, horse and rider are walking, that is, they are not yet in a position to canter (the task is to transition to canter via trot, see line 1). While they are still walking, the instructor explains the general aids for canter (inside leg on outside leg back), which are to be performed during the upcoming transition to canter but not right now (the trot has yet to be ‘established’, see line 1). While the instructor delivers them, horse and rider start trotting (Image 1), which means they will soon be ready to canter. After a 1.5s
silence, the instructor gives the local NOW directives eyes up, which steps up from 184 to 287 Hz (Fig. 10, lefthand margin), and and then sit up, which reaches 367 Hz (righthand margin; no pitch trace can be picked up for sit). The two preparatory directives target immediate, locally relevant adjustments to the rider’s position, which apply specifically at this moment. As the second directive ends, horse and rider transition to canter, that is, they perform the instructed action. After a 0.7s silence, the successful compliance is receipted with good, which rises to 522 Hz and then falls to 200 Hz, close to the starting point of the three-part structure. The following well done repeats the high step-up and low fall with a reduced pitch span. Images 2-5 show horse and rider during their preparation and achievement of the transition from trot to canter.

![Pitch trace](image)

Fig 10. Pitch trace, Extract (5), lines 3-4.

Like the prosodic escalation on sit kick canter in Extract (2), this case shows gradually escalating pitch and loudness with the preparation and eventual accomplishment of an
instructed activity. Both instances represent a form of diagrammatic iconicity\(^3\) in that their prosodic escalation on minimal syntactic structures mirror the progression and culmination of a series of instructed actions.

As Extracts (1) – (5) show, instructions to perform mobile tasks can be accompanied by a stepwise rise in pitch register and often also in loudness, and occasionally by changes in voice quality. A final pitch peak is typically followed by a fall and a return to less marked prosody. Unlike previously described prosodic patterns that depict the structure, rhythm, or synchrony of an instructed activity (Keevallik 2020; Ehmer 2023; Hofstetter & Keevallik 2023), escalating prosody is used as a resource for vocally depicting recipients’ physical effort, mobility, and progression. The progression often ends in a final peak, physically and prosodically. The intensified format, which requires its speaker to exert additional breath force, mirrors the embodied effort the recipient is being asked to invest. Through escalating prosody, instructors co-design and mobilize the preparation, accomplishment, and closing of learners’ instructed activities. Instructors’ escalating talk can cover the full set of component actions that horses and riders must perform to accomplish the instructed activity; or it can target individual actions within larger activities, or series of repeated actions. Escalating prosody thus provides a temporal frame; however, in the present data the temporal co-construction is not one of pressure and urgency due to a delayed compliance with instructions, as shown previously by Mondada (2017) and Okada (2023). Instead, the gradual build-up depicts mobile progression and sustained physical effort.

6.2 Unfolding praise

---

\(^3\) I am grateful to one of the anonymous reviewers for suggesting this concept here.
A specific environment for escalating prosody occurs when instructors give repeated positive assessments as the instructed activity unfolds over time, as seen in Extract (1). This is the case in six of the 25 instances in the corpus. All six involve either the same or similar lexical items. The repeat assessment tokens are not ‘multiple sayings’ (Stivers, 2004) but instead assess the same activity at different moments of progression. In the following example, horse and rider have been asked to perform ‘shoulder in’, a movement where the horse moves down the long side of the arena with their shoulder turned slightly to the inside.

(6) 20220605 Lesson 1 3:8.25

1 I: so you’re going to imagine,  
H/R: >>walk-->>  
H/R: >>into and around corner-->>

2 I: that you’re going on to your diagonal line,=

3 I: =and then change your mind at the last minute;=

4 I: =and ride straight.--+(0.6)  
H/R: ------------------------+ straight line-->>

5 I: your shoulders are+ pointing that way, #(0.6)  
H/R: ------------------------+...

Image #Image1

6 I: but your hips+ are pointing towards me.==#==there you go;#  
H/R: .................+ shoulder in-->>

Images #Image2 #Image3

7 --> I: <<h> there you go;# (0.3) <<hh> there you go;#
The extract shows the second time horse and rider are performing the shoulder in during this lesson. At lines 1-6, the instructor first tells the rider how to conceptualize the riding of the shoulder in and then how to perform it in her body.4 Once horse and rider are performing the instructed activity, the instructor begins to praise them and continues doing so until she gives the directive to finish (lines 6-8). On its first occurrence, the phrase there you go is delivered in a pitch range similar to prior talk. The next two iterations gradually step up in pitch. The stressed syllable go rises from a peak at 164 Hz on its first occurrence (line 6) to peaking at 284 Hz on the second iteration and at 322 Hz on the third iteration. The final assessment that’s it peaks at 428 Hz. As the instructor closes the exercise, her pitch drops to approximately 215 Hz (Fig. 11, righthand margin). The escalating feature in this extract is pitch alone, as there is no noticeable change in loudness or voice quality. The instructor’s first

---

4 See Szczepek Reed (2021) for a distinction between conceptualized and embodied learnables.
two iterations synchronize with the horse’s walk, but the third iteration is delayed. As in previous examples, repeated lexical items are prosodically escalated in a mirrored representation of the unfolding to-be-sustained activity.

Fig 11. Pitch trace, Extract (6), lines 6-9.

An example from the German collection also shows escalating prosody with a build-up of praise. Horse and rider are practicing ‘leg yield’, that is, a movement where the horse is asked to step sideways. They have just leg-yielded to the left and are now being asked to go straight and then leg-yield to the right.

(7) FOLK_E_00491 19.06

1   I:  GRADe? und UMstellen. *(0.2) || || (1.2)
     straight? and the other way. || ||
     H/R:  *leg yield to right>*

2   I:  jetzt mit’m LINKen schenkel.=
     now with your left thigh.

3   -> I:  =<<ch> J:A: (0.2) ↑GU::T; (0.2) ↑↑GU::T; (0.2) ↑↑↑GU::T;*
     ↑'e::s. ↑goo::d; ↑↑goo::d; ↑↑↑goo::d;
     H/R:  ---------------------------------------*

4   -> I:  *+(1.0) ↑JA: BESSer.>
     ↑↑ye:s better.
When horse and rider start leg-yielding to the right, the instructor praises their successful performance, as the horse maintains his sideways movement for as long as the rider is giving him the aids to do so. The first of the positive assessment tokens (*ja*) already steps up to a higher pitch than previous talk, peaking at 444 Hz. It is delivered with a rising-falling pitch contour and lengthening on the vowel. As Figure 12 shows, each subsequent token steps up higher than its predecessor, repeating the rising-falling contour and the vowel lengthening. The final peak reaches 617 Hz. The final phrase *ja besser*, which follows a 1sec pause, repeats the contour and the lengthening, and maintains a high pitch register, peaking at 515 Hz. This last phrase occurs once horse and rider have completed the task, while the earlier, gradually rising tokens accompany the performance as it emerges. There is no clear rhythmic alignment between the instructor’s escalating praise and the horse’s strides. The prosodically escalating assessments construct the horse’s and rider’s activity as one that unfolds successfully, while highlighting the need to sustain the physical effort. That the rider shares the appreciative stance is evident from her patting the horse’s neck immediately afterwards (line 4).
Unlike directive giving, which is action-initiating, praise is a responsive action. The near-simultaneous delivery of praise with the successfully performed actions resembles the ‘sequentially ordered simultaneities’ identified by Mondada (2017: 91), who shows that the verbal shaping of emerging actions results in a parallel, rather than a sequential ordering of instructions and their responses (see also Keevallik 2020: 170). A related finding is reported in Okada (2018), who shows that boxing coaches can give directives before, during, and even after the boxer has performed the instructed action. In giving praise as the requested activity unfolds, instructors use the affordances of concurrent prosodic escalation to, firstly, emphasize the ongoing intense nature of the activity and praise the very sustaining of the activity as such, and secondly, to choose explicit support for the horse’s and rider’s actions over further instructional intervention.
6.3 Escalating corrections

In three instances, escalating prosody is found with correction talk. This typically involves repeated directives in response to learners’ lack of, or unsuccessful compliance. In contrast to Mondada’s (2017) analysis of quickly delivered directives in the face of delayed compliance by the driving pupil, the cases in this corpus are not produced with fast speech rate. All three cases come from the German, mobile phone-assisted collection; and all three involve lexical repetition, again addressing moments of the activity as it unfolds. In the following extract, horse and rider are trotting around the arena, practicing the horse’s bending around the corners. As they come out of one corner, they start to veer off the track and towards the middle of the arena.

(8) FOLK_E_00491 6.57

1 I:  MEHR durch die Ecke; * ne? MEHR die bIE**gung. (0.6)
    more through the corner; yeah? more bend.
    H/R:  >>trot through corner* *trot straight---**veer to left->*

2 -> I:  ↑nIch* *:Abdriften- <<ch> ↑↑nIch *:Abdriften- <<hh> ↑↑↑nIch
        ⦿
        ↑don’t drift off; ↑↑don’t drift off; ↑↑↑don’t drift off;
        H/R:  ----- *gradually return to track->*

3 I:  (0.2) || (1.2)* *(0.9) NE? sei da STRENGer mit dem inneren
        ||
        yeah? be stricter with your inside
        H/R:  --------------* *back on track->>

4 I:  schenkel. jetz AUCH wieder. mEhr in die ECKe. KOMM.
        thigh. and there again. more into the corner. come on.

As horse and rider start to drift to the inside, the instructor gives the correction nich abdriften (‘don’t drift off’) three times. The German construction is a deontic infinitive, which Deppermann (2012) shows can perform, amongst other things, the action of prompt or command (‘Aufforderung/ Gebot’, p. 246). The first occurrence begins with a pitch step-up on nich, followed by a step-down on ab-. abdriften is delivered with level pitch. This contour
is repeated twice with gradually rising pitch register. *nich*, the onset syllable that performs the step-up on each occasion, rises from 306 Hz to 364 Hz and finally to 407 Hz (Figure 13, righthand margin). The instructor’s turn contains only pitch escalation, without an increase in amplitude. The corrective talk accompanies the horse’s and rider’s attempts at steering back towards the track, which is also supported with a click from the instructor. She resumes her instruction talk once they have reached the track with a directive to the rider to be ‘stricter’ with their inside leg (lines 3-4).

When escalating prosody is being used with corrections, instructors orient to moments where riders need specific support because something is going wrong. In these instances, escalating prosody does not accompany an activity from beginning to end but instead targets certain components of the activity. This finding resembles Okada’s (2023) evidence of repeated
imperatives by boxing coaches that increase in speech rate, loudness, and pitch when directives are not complied with straight away. The lexical repetition of the to-be-corrected component provides a temporal timeframe for the duration of the correctable activity and co-constructs the corrective activity in real time.

7. Discussion

Recent work has shown that prosody can synchronize others’ bodies (Keevallik 2020; Hofstetter & Keevallik 2023), enact their sensations (Keevallik et al. 2023), and help shape others’ actions as they unfold (Mondada 2017; Simone & Galatolo 2021). These studies have revealed that sound patterns are an in situ resource for understanding and relating to the embodied actions of others. The findings of the present research show that in addition, prosody also manages the mobile progressivity and embodied intensity of others’ actions. It does so by offering speakers flexibility in terms of time (duration, tempo, rhythm) and breath force (pitch, amplitude, phonation), providing depictive affordances that are unique among the linguistic resources. This flexibility, paired with minimal syntax (single words, short phrases), is ideally suited to depict and co-design transient mobile actions as they unfold through time. Sound patterns that have previously been conceptualized as representations of affective involvement were shown to be used for instructional, rather than emotional purposes: instead of displaying a speaker’s own stance, such as enthusiasm during football commentary, here they are used to bring about responses in others. Rather than displaying personal involvement, they mirror the embodied energy that is requested of recipients, depicting vocally the sustained, progressing, and intensifying effort they seek to elicit. The activities that are co-constructed through escalating prosody often develop towards a peak, and escalating prosody is a resource for mobilizing others to reach and maintain such peaks.
While directives that are delivered with prosodic escalation are always local NOW directives, to be performed immediately, they are not linked to ‘brevity, speed, and urgency’ (Mondada 2017, pp. 71), that is, they are not done to elicit faster or earlier performance. Instead, they depict and guide to-be-sustained activities as they unfold.

This study included two recordings where the participants used mobile phone technology, with the instructor sitting on a chair by the side of the arena, with the mobile phone on her lap, and the rider participating through earphones. These recordings show a noticeable difference in overall loudness compared to the riding lessons recorded without mobile phones, as the instructor in the technology-assisted lessons speaks more quietly throughout. In addition to the technological affordances, which mean she does not have to project her voice across a physical distance, the instructor’s seated position is likely to affect her vocal production. Nevertheless, these recordings contain instances of escalating prosody (8 cases), which involve mostly stepwise increases in pitch register. There are no escalations of amplitude and no harsh voice quality. However, the rest of the collection also contains instances of pitch-only escalation, as in Extract (7). It is likely that conditions of the physical world have an impact on how speakers vocally depict bodily force and effort, alongside many other factors. Different modes of embodied skills instruction present an important opportunity for future research.

The finding that prosody can co-construct progressivity and physical effort is unlikely to be limited to horse-riding lessons. Keevallik et al. (2023) show that there are diverse contexts for giving sound to others’ bodily sensations; the same is likely to be true for depicting mobility, effort, and progression. Future work will show the affordances of escalating prosody for other communicative contexts, instructional or otherwise.
References


Imperative turns at talk: the design of directives in action (pp. 65–101). Amsterdam: Benjamins.


Appendix

Verbal transcription, adapted from Selting et al. (2009/2011)

(.) micro pause

(1.2) measured pause

::: lengthening

ACcent primary pitch accent

Accent secondary pitch accent

<<h>> high pitch register

<<f>> forte

.h hh .hhh in-breath

h hh hhh out-breath

l click

[talk

[talk overlap

Phrase-final pitch movements:

? rise-to-high

, rise-to-mid

- level

; fall-to-mid
fall-to-low