Computers and Electronics in Agriculture xxx (2011) xxx–xxx

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

Acoustic monitoring system to quantify ingestive behavior of free-grazing cattle

William M. Clapham[a,*], James M. Fedders[a], Kim Beeman[b], James P.S. Neel[a]

[a] USDA-ARS, Appalachian Farming Systems Research Center, 1224 Airport Rd, Beaver, WV 25813, USA
[b] Engineering Design, 262 Grizzly Peak Blvd, Berkeley, CA 94708, USA

ARTICLE INFO

Article history:
Received 6 May 2010
Received in revised form
23 December 2010
Accepted 14 January 2011

Keywords:
Acoustic monitoring
Intake
Forage
Cattle
Grazing behavior

ABSTRACT

Estimating forage intake by free-grazing livestock is difficult and expensive. Previous approaches include behavioral observation, ratio techniques using indigestible markers, mechanical recording of ingestive jaw motion, and acoustic recording of ingestive behaviors. Acoustic recording shows great potential but has been limited by the difficulty and time required to manually identify and classify ingestive events. We developed an acoustic recording and analysis system that automatically detects, classifies, and quantifies ingestive events in free-grazing beef cattle. The system utilizes a wide-frequency acoustic microphone close to the animal's mouth, mathematical signal analysis to detect and measure ingestive events, and streaming data analysis that can handle an unlimited amount of data. Analysis parameters can be reconfigured for different animals, forages and other changing conditions. The system measures the acoustic parameters of ingestive events, such as duration, amplitude, spectrum and energy, which can support further event classification and become the inputs to a forage intake model. We validated our detection and classification technique against the results of trained human observers based on field studies with grazing heifers. The software detected 95% of manually identified bites in an event-by-event comparison. Field observations and sound attenuation analysis indicate that sounds from adjacent livestock and ambient pastoral environments have an insignificant effect upon the integrity of the recorded acoustic data set. Wideband acoustic analysis allows us to identify ingestive events accurately and automatically over extended periods of time.

1. Introduction

Sound metrics, including frequency and amplitude, can be used to classify and quantify food and ingestive processes. Acoustic analysis was used to quantify texture (crispness/crunchiness) in food. Liu and Tan (1999) studied snack food crispness and demonstrated that sound features corresponded ($R^2 = 0.89$) with a trained sensory panel, concluding that sound signal analysis provided an effective measure of crispness. Similar data were collected measuring apple and potato crispness (Zdunek and Bednarczyk, 2006). Acoustic envelope detectors were developed (e.g. Stable Micro Systems, Surrey, UK) to quantify the crispness and sensory qualities of biscuits and other fresh and processed foods.

Forage intake by grazing livestock is one of the keys to understanding forage grazing system dynamics (Ungar, 1996). However, estimating intake of free-ranging livestock is difficult and expensive. Technology and improved methods have significantly improved our ability to collect grazing behavior data. Procedures to estimate intake include indirect methods such as ratio or index techniques, where intake is calculated via measures of digestibility (Cordova et al., 1978), and direct methods such as direct behavioral observation; mechanical recording of chews, bites, and jaw activity using jaw sensors (Chambers et al., 1981; Champion et al., 1998); acoustic recordings in combination with video recordings or direct observation (Griffiths et al., 2006; Laca et al., 1992). The development of jaw sensors and small data recorders (Rutter et al., 1997) provided a wealth of data regarding ingestive behavior, particularly because software to classify the data was developed to quantify jaw movement events (Rutter, 1998). However, estimates of intake require calibration of the relationships between bite count and forage ingested and modeling variation in bite size. Some success was achieved by combining video and acoustic recordings of ingestive behavior combined with short-term studies of mass difference from 0.14 m^2 field-grown, sods placed in metal trays (Laca and WallisDeVries, 2000). Acoustic methods pioneered by Laca et al. (1992, 1994), and used by Galli et al. (2006) and Ungar and Rutter (2006) utilized "an inward-facing microphone mounted on the forehead of the animal" to record the sounds of bites and chews. Ungar and Rutter (2006) demonstrated that data collected using an inward-facing microphone corresponded to data collected using the IGER Behaviour Recorder in 10-min grazing sessions using

* Corresponding author. Tel.: +1 304 256 2857; fax: +1 304 256 2921.
E-mail addresses: William.Clapham@ars.usda.gov (W.M. Clapham), jim.fedders@ars.usda.gov (J.M. Fedders).

0168-1699/$ – see front matter © 2011 Published by Elsevier B.V.
doi:10.1016/j.compag.2011.01.009

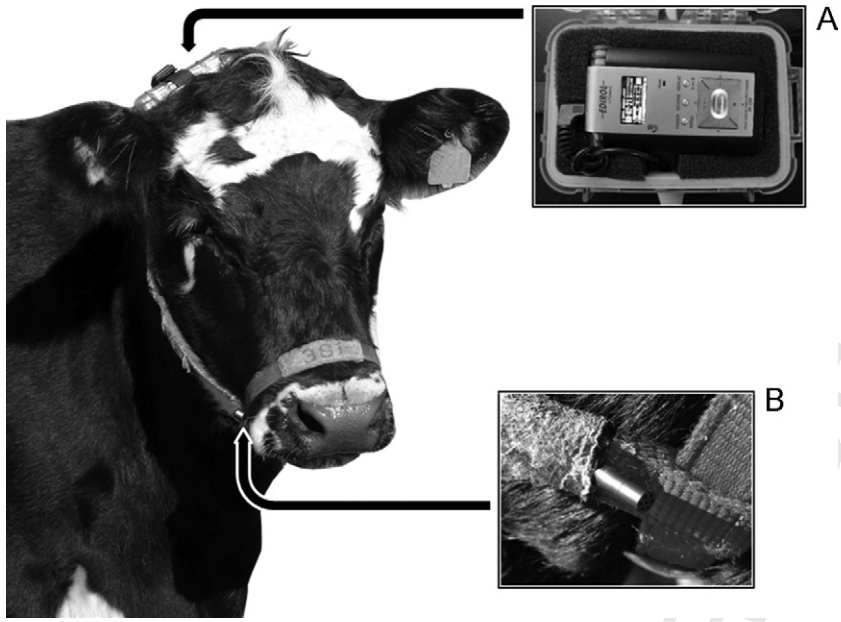

Fig. 1. Photograph of a heifer wearing a halter with attached digital recorder and microphone. Inset shows details of recorder in the protective plastic case.

six cattle. Although acoustic methods demonstrate great promise for recording and quantifying ingestive events, manual classification of these events is difficult and time consuming and in need of automation (Ungar and Rutter, 2006). Milone et al. (2009) created software that used hidden Markov models to automate the identification and classification of ingestive events in sheep and classified bites and chews with an accuracy of 58 and 89%, respectively.

In this paper, we describe the development of a digital audio recording and automated event classification system that records grazing sounds, detects bite events and compiles grazing event data (bite number and acoustic event parameters). The objectives of this report are to: (1) describe the hardware and software components and processing steps; (2) compare the spectral characteristics of ingestive events recorded over wide and narrow frequency ranges, to demonstrate the need for wide-frequency acoustic data for accurate automated detection of bite events; (3) establish the acoustic features required to differentiate and classify ingestive events; (4) document the amount of acoustic cross contamination from animals grazing nearby; and (5) use manual analysis of audio–video recordings to validate the ability of the automated system to detect and classify bite events.

2. Methods and materials

2.1. Field conditions

Ingestive behavior was investigated at West Virginia University Willow Bend Farm near Union, WV, USA (37.547°N latitude, 80.528°W longitude). Halter-trained, 16–18 month old, angus-cross steers or heifers (450–550 kg live weight) were used during the experiments. The free-ranging animals were maintained on mixed, perennial pasture consisting primarily of tall fescue (*Festuca arundinacea* Schreb.), orchardgrass (*Dactylis glomerata* L.), bluegrass (*Poa pratensis* L.) and white clover (*Trifolium repens* L.). Recording sessions were conducted between the hours of 8:00 AM and 1:00 PM local time between July and October over five years. During a recording session, the animals were given access to either mixed perennial pasture, alfalfa–orchardgrass pasture or a pure stand of triticale (X *Triticosecale* Wittmack) (a mixture of Trical 2700 and Trical 336; Resource Seeds Inc. P.O. Box 1319, Gilroy, CA 95021) that had been established in early August.

2.2. Hardware components and setup

The recording system (Fig. 1) was designed to have minimal intrusion on the behavior of the livestock. The system consisted of a digital recorder (Edirol R-09 24-bit recorder, Program Version 1.20, Roland Corporation US, 5100 S. Eastern Ave., Los Angeles, CA 90040-2938) and omni-directional lavalier microphone (Sennheiser ME 2-US, Sennheiser Electronic GmbH & Co. KG, 30900 Wedemark, Germany) mounted on a 1-inch nylon cow halter (Weaver Leather, 7540 CR 201, PO Box 68, Mt. Hope, OH 44660). The recorder was placed inside a water resistant plastic enclosure (Pelican 1020 Micro Case, Pelican Products, Inc., 23215 Early Avenue, Torrance, CA 90505) and bolted onto the back strap of the halter to ride behind the head of the animal. The microphone was attached to the front strap of the halter 5 cm from the right corner of the animal's mouth. Four-inch wide Vetrap tape (3 M Animal Care Products, St. Paul, MN 55144-1000) was used to secure the microphone and microphone cable to the halter.

Sound data was recorded onto a 4 GB SD memory card (Sandisk Extreme III SDHC Card Sandisk Corporation, 601 McCarthy Blvd., Milpitas, CA 95035) in the Edirol R-09. All recordings were made at 44.1 kHz sampling rate and 16-bit resolution, providing a nominal 22 kHz recording bandwidth and 96 dB dynamic range, and stored in the WAV (Waveform Audio) file format. Recorded sound files contain the voltage output from the microphone, representing the time-varying acoustic pressure at the microphone diaphragm. Voltage values can be converted to numerical sound pressure by applying a calibration factor incorporating microphone transducer gain ($\mathrm{V\,Pa^{-1}}$) and amplifier gain. Prior to each recording session, the recorder input level, sampling rate and bit resolution were set; the recorder was secured inside the plastic enclosure; and the halter was secured on the animal. Four to six animals grazed together during each recording session in paddocks that were approximately 0.1 ha in size.

2.3. Sound file processing and analysis

Files from each recording session were uploaded onto a Dell Optiplex 745 personal computer (Dell Inc., One Dell Way, Round Rock, TX 78682, USA) (3.40 GHz Intel Pentium D CPU; 4 GB RAM; Microsoft Windows XP Professional, version

Please cite this article in press as: Clapham, W.M., et al., Acoustic monitoring system to quantify ingestive behavior of free-grazing cattle. Comput. Electron. Agric. (2011), doi:10.1016/j.compag.2011.01.009

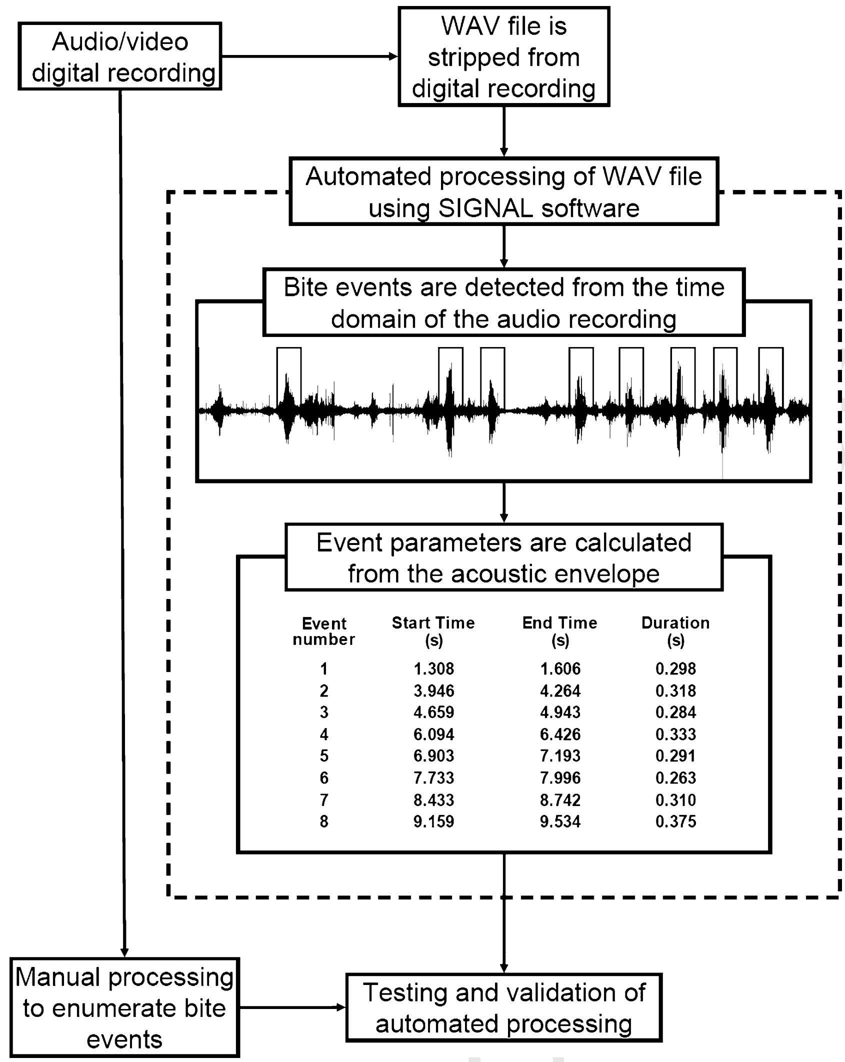

Fig. 2. Schematic of testing and validation procedure for automated processing of audio recordings. The waveform represents acoustic pressure over 10 s, recorded from steer 751 while grazing mixed pasture on July 28, 2005. Rectangular boxes on the waveform mark bite events detected by the SIGNAL program. Measured event parameters are shown in the box below the waveform, as an example of program output.

5.1.2600). Audacity software for Windows (version 1.3.5 beta, http://audacity.sourceforge.net/) was used to prepare the raw WAV files for analysis. The stereo files created by the R-09 recorder were reduced to monaural files by extracting one channel. A high-pass filter (rolloff = 24 dB, filter quality = 0.1, cutoff frequency = 600 Hz) was applied to reduce wind sounds and other low frequency noise. In future work, we will attempt to eliminate or reduce the need for this filter by improving microphone wind-resistance.

Identification, enumeration and measurement of bite events in the pre-processed files were performed using the SIGNAL sound analysis program for Windows (version 5.00.28, Engineering Design, 262 Grizzly Peak Blvd, Berkeley, CA 94708, USA, www.engdes.com). SIGNAL analyzed each sound file and automatically detected and measured bite events, recording the measurement data into a log file.

The SIGNAL software processed the monaural WAV files at approximately 10 times real–time, i.e., analyzing 10 min of acoustic data per minute. The software operates in a two-step process (Fig. 2). First an event is detected, then event parameters are measured and recorded in a log file. This process is performed repeatedly, from the beginning of the file to the end with the goal of detecting every target event in the file. SIGNAL detects events based on sound characteristics such as frequency, intensity, duration and time between events (Table 1). The values we chose for these parameters collectively define a bite event to the software and enable it to detect bite events. Initial detection values were selected based on the differing amplitude and frequency characteristics of bite vs. non-bite events. Detection values were then refined through trial and error by comparing automated and manual bite counts.

Table 1
Parameter settings used to detect bites from acoustic recordings of grazing sessions with SIGNAL software.

Parameter	Value
Low frequency cutoff	17 kHz
High frequency cutoff	None
Envelope decay time	15 ms
Detection threshold	0.013 V
Minimum event gap	250 ms
Minimum pulse length	1 ms
Minimum event length	100 ms
Maximum event length	1000 ms
Pre-event time extension	100 ms
Post-event time extension	100 ms

Please cite this article in press as: Clapham, W.M., et al., Acoustic monitoring system to quantify ingestive behavior of free-grazing cattle. Comput. Electron. Agric. (2011), doi:10.1016/j.compag.2011.01.009

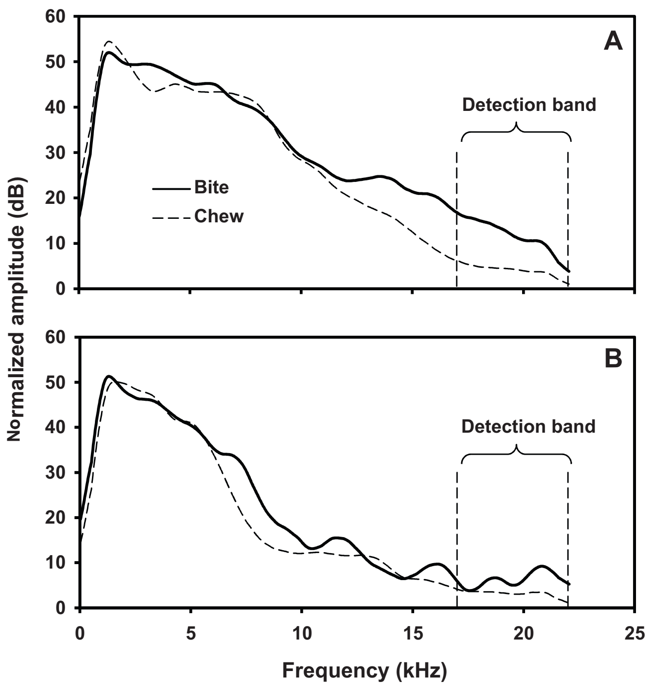

Fig. 3. Spectral density plots of: (a) wideband (0–20 kHz) recordings and (b) narrowband (0–8 kHz) recordings of one bite and one chew event of a heifer grazing vegetative triticale. Spectral energy below 600 Hz was removed from all signals by a high-pass filter during preprocessing. Spectra were derived from a 16,384-point Fourier transform, adjusted for 1 Hz spectral bandwidth and smoothed with a 1000-Hz rectangular window. Co-plotted spectra are normalized for equal RMS power to contrast spectral distribution. Brackets indicate the 17–22 kHz frequency band used to detect bite events.

Key distinguishing characteristics of bite events, relative to chewing or other sounds, are the high frequencies produced by the initial shearing or ripping of forage. Therefore we programmed the detection software to evaluate only energy at frequencies between 17 kHz and the upper recording limit of 22 kHz (see Table 1), as shown in Fig. 3. For purposes of bite detection, we programmed the software to register an event start time (T_s) when event amplitude in the 17–22 kHz range exceeded a given detection threshold and an event end time (T_e) when amplitude subsequently dropped below this threshold. Events were discarded as spurious if, for example, their durations were shorter than the minimum designated event length or longer than the maximum designated event length.

Detailed examination of time-domain representations of bites indicates a low level of bite energy immediately prior to T_s as sound energy increases from background levels to the detection threshold and immediately after T_e as bite energy dissipates and drops below the threshold. We programmed SIGNAL to include this energy by using an event measurement period of $T_s - 100$ ms to $T_e + 100$ ms as indicated by the pre-event and post-event time extensions in Table 1. We expect our technique will detect and extract the bite portion of "chew–bite" events (Laca and WallisDeVries, 2000) consisting of a chew followed immediately by a bite. We expect contamination due to spurious inclusion of the chew segment will be small because: (1) the 600 Hz high pass filter utilized prior to event detection removes much of the chew energy and (2) the 100 ms time extension matches the small separation depicted between the chew and bite segments of the chew–bite event illustrated in Laca and WallisDeVries (2000).

Our approach to bite detection is not intended to measure bite duration precisely but rather to detect automatically the occurrence of bite events with high reliability, count bite events and measure the sound energy produced when forage is sheared in each bite. For example, our approach to bite detection does not include the time taken by the animal to gather forage with the tongue and bring it into the mouth prior to shearing, a process that will vary with sward structure and composition.

This phase of our work did not require calculating the absolute energy of acoustic events. However, we did compare relative acoustic energy levels, for example, to estimate the percentage contamination of bite sounds by adjacent animals (Section 2.5.1). For this purpose, we calculated the total energy flux of an event, defined as $\int \mathrm{J\,m^{-2}\,s^{-1}}\,dt$. Since instantaneous energy flux in a plane acoustic wave is $p^2/\rho_0 c$, where p is pressure and ρ_0 and c are, respectively, the density and propagation velocity of the medium, total energy flux depends on $\int p^2\,dt$. Since the amplitude of our acoustic data is proportional to pressure, our program calculated the time-integrated squared amplitude for each bite as a measure of relative event energy.

2.4. Comparison of wideband and narrowband recordings of ingestive sounds

Our technique for automating detection and classification of bite events distinguishes bites from chews based on high-frequency (17 kHz and above) characteristics and therefore requires full bandwidth acoustic recordings. To confirm this, we made "wideband" (0–22 kHz) and "narrowband" (0–8 kHz) recordings of the same ingestive events. We use these terms to refer to these bandwidths throughout this paper. Wideband recordings were made with a halter-mounted ME 2 acoustic microphone attached near the animal's mouth and narrowband recordings were made with a forehead-mounted piezoelectric microphone fashioned from a 2.5 cm diameter piezoelectric transducer (Edmunds Scientifics, Tonawanda, NY 14150, USA). Signals were recorded simultaneously on separate channels using the stereo capability of the Edirol R-09.

This system was mounted on one heifer grazing triticale on October 20, 2009 and on another heifer grazing triticale on October 22, 2009. Five bite events and five chew events were randomly selected from wideband data and similarly from narrowband data for a total of 20 exemplars, to avoid crowding on the PCA plot. Two temporally synchronized monaural files, one wideband and one narrowband, were created from the stereo file for each event using Audacity software. The frequency spectra of the 40 files (2 animals $\times$ 10 events animal^{-1} $\times$ 2 files event^{-1}) were analyzed by SIGNAL using Fourier transform techniques. For each file, relative spectral amplitude in dB was determined at 86.1 Hz intervals across the spectral range from 0 to 22000 Hz. These 256 values for each of the 40 events were then subjected to principal component analysis using the PRINCOMP procedure of SAS for Windows, version 9.2 (SAS Institute, Cary, North Carolina 27513, USA).

2.5. Estimating acoustic contamination of recordings

Our recordings of acoustic bite events can be contaminated in two ways: by bite sounds from other animals and by non-target noise events such as insects or jet plane flyovers.

2.5.1. Cross-contamination from other bite sounds

Our studies involved multiple animals grazing together in close proximity, which creates the possibility that a recorded bite from one animal (the target) may include bite sounds from nearby (non-target) animals. We call this bite-sound cross-contamination. Significant cross-contamination can degrade the automated detection process with false triggers, as well as corrupt quantitative measurements of detected bite events, and has been noted as a serious concern (Ungar and Rutter, 2006). We quantify cross-contamination as the fraction of recorded target bites that contain significant energy from the bite sounds of non-target animals, defined as a contaminating energy level of 1% or greater relative to the target bite sound energy, as measured at the target microphone. The 1% contamination level was selected as the threshold below which contamination would have minimal impact on bite energy measurements.

We could not directly measure bite sound cross-contamination in our field recorded sound signals as a function of animal proximity due to the difficulty of obtaining time-correlated acoustic and proximity data without suitably tame and trained animals. Instead we modeled cross-contamination in the following way. We observed inter-animal separation distances under field conditions, measured the average rate of bite production in field recordings, and applied the physics of sound attenuation in air to calculate contaminating bite energy at varying distances. The equation for acoustic radiation in free space states that energy attenuates in proportion to the squared distance from the sound source. Assuming target and non-target bites have similar source energy, the sound energy of the non-target bite will exceed 1% of target bite energy when the non-target animal is less than 10 times as far from the recording microphone as the target animal's mouth. Since the recording microphone is mounted 5 cm from the target's mouth, we are concerned with animal encounters closer than 50 cm, in which the contaminating acoustic energy would be 1% $(5\,\text{cm}/50\,\text{cm})^2$ or more of the target energy.

We observed six heifers while they grazed together within a 46 m × 34 m paddock of vegetative triticale at 20-s intervals over five periods of 5–10 min each. At each interval, we counted the number of animals whose heads were within 1 m of each other, as this separation distance was easier to estimate in the field than the 0.5 m critical distance. The number of interactions at 0.5 m or less is estimated by a linear interpolation between 0 and 1 m.

2.5.2. Contamination from non-target sounds

Non-target noise events include intermittent sounds, such as flies, birds, animal vocalizations, aircraft, farm equipment and road traffic, and continuous sounds, such as crickets and grasshoppers. As with cross-contamination, our goal was to estimate the contamination of measured acoustic energy in the target bite. Intermittent sounds – such as aircraft flyovers – are short-duration, potentially high-intensity, and usually infrequent. We estimated the statistics of this contamination in terms of the fraction of recorded bite events that would be affected. Continuous sounds – such as insects – are long duration and low intensity. For example, large populations of crickets in our pastures during mid to late summer create continuous background sounds that are present in every event in the data set. For these we estimated the ratio of target to non-target acoustic energy and from this ratio we calculated the spurious increase in target energy as a percentage error. We analyzed two representative 10-s sound samples, each containing a bite sequence with background cricket sounds and a segment of cricket sounds without bites. Data was high-pass filtered at 600 Hz to remove wind noise and other low-frequency ambient sounds. We calculated the ratio of bite energy (energy of bite with contaminating cricket sounds minus energy of cricket-only sounds) to cricket energy for each sample segment.

2.6. Calibration and validation of automated bite detection

Audio/video recordings of grazing activity were used to calibrate and validate bite detection parameters used by the SIGNAL software (Fig. 2). Digital camcorders were used to record ingestive behavior of three animal subjects. Two animals (steers 751 and 527) were recorded grazing mixed perennial pasture on July 28, 2005 and one animal (steer 710) was recorded grazing alfalfa on September 8, 2005. Continuous 15–30 min recordings of each animal on each date were made with Canon Elura 85 Digital Camcorders using Maxell Mini DV Digital Video Cassette tapes. Audio was transmitted from the halter-mounted acoustic microphone to a camcorder using a Samson AL1 UHF transmitter mounted on the halter behind the neck of the animal. Camera operators were stationed outside of the paddocks where the animals grazed. When multiple animals were recorded simultaneously on July 28, 2005, the transmitters were set to different frequencies to isolate transmission of audio from each animal to separate cameras.

The single audio/video recording from each animal and date was divided into 1–5 min segments for analysis and converted to MOV files using iMovie software (Apple Computer, Cupertino, CA 95014). We created three files representing 15 min of data for steer 751, four files representing 20 min of data for steer 527, and five files representing 15 min of data for steer 710. The number of bites recorded on each MOV file was manually tallied by a trained observer while reviewing the combined audio and video tracks. Manual classification was based on synchronized CD-quality audio and close-up video that provided visual details of the distinctive mouth and head movements associated with bite events. To estimate the accuracy of our counts, we repeated them using a second trained observer. The audio track was extracted from each MOV file using iMovie to create digital audio WAV files (44.1 Hz, 16-bit, monaural) for automated bite analysis using the SIGNAL software.

For each animal and date, one WAV file was chosen at random as a calibration file for the SIGNAL program. SIGNAL detection parameters (Table 1) were adjusted until the SIGNAL-derived bite count was within 2% of the manual bite count for the calibration file. Generally, calibration involved minor adjustment to the detection threshold level among animals grazing the same forage type and larger adjustments to the threshold between forage types. Other detection parameters generally did not change. Once calibrated, the SIGNAL program was then used to count the bites from the remaining files for that animal and date without any further parameter adjustments. In this manner, SIGNAL-derived bite counts were determined for all of the WAV files.

Automated detection was validated against the manual baseline in two ways. First, automated and manual bite counts were compared. The SAS GLM procedure was used to test for significant differences between the manual and SIGNAL-derived bite counts. The model was a repeated-measures ANOVA with between subjects factors. Differences between the two count methods were evaluated by assessing differences in bite count within recordings and interaction was evaluated to assess any differences in count method among the animals. We also calculated the standard deviation of the residual error of the automated bite counts compared to the manual bite counts.

Second, automated and manual bite sequences were compared event by event. One 5-min WAV file was selected at random for this detailed analysis. Using the audio and video recording, a trained observer recorded the mid-bite time coordinate of every bite event in the file (blind to the automated result on that file). Manual and automated bites were then compared one by one. Bites were considered matched if the manually derived bite time fell between the start and end times of an automatically detected bite. This analysis produced three counts: matched bites, false positives (a non-bite sound detected as a bite by the automated system) and false negatives (a manually identified bite missed by the automated system).

3. Results

The self-contained, halter-mounted recording system was lightweight and did not appear to restrict animal activity. Direct observation suggested that the animals exhibited normal grazing behavior while wearing the halters. The animals were typically eager to graze fresh forage during the experimental trials after

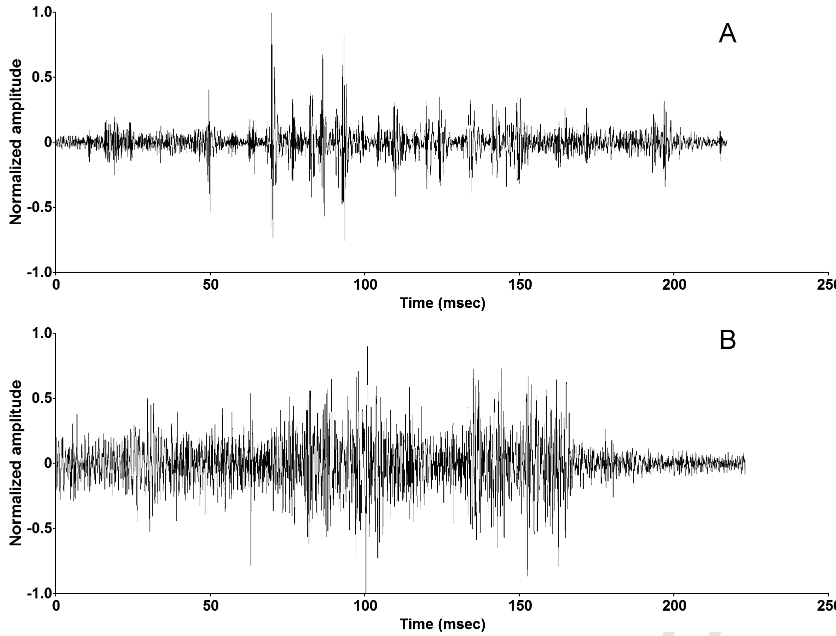

Fig. 4. Wideband acoustic recordings of: (a) the bite event and (b) the chew event shown in Fig. 3a. Signals are displayed as time-domain waveforms and are normalized to unit peak amplitude.

spending the previous night penned with a limited amount of forage and/or dry hay. Laboratory tests showed that the 4 GB SD memory cards could hold up to 6 h and 24 m of sound recordings, longer than any of the trials conducted thus far. A pair of fully recharged batteries powered the recorders long enough to fill the SD cards in laboratory tests.

Sound signal data can be expressed as sound intensity vs. time (Fig. 4) or as sound intensity vs. frequency for a given time period (Fig. 3). Both representations provide insight into the recorded sounds. A typical bite generated sound for a duration of approximately 0.1 s (Fig. 4a) and the sound spanned a wide frequency range (Fig. 3a). Frequencies below 600 Hz are excluded by the high-pass filter applied during pre-processing. Amplitude declines between 8 kHz and 22 kHz, the upper limit of our recording system (Fig. 3a), but that range is important for detecting and classifying bite signals.

3.1. Importance of wideband acoustic recordings

We performed a principal component analysis (PCA) on the spectra of wideband and narrowband recordings of the same bite and chew events from the two animals under study (Fig. 5). The first two principal components accounted for 96% of the variation in the spectral signatures. In Fig. 5a, bites and chews are effectively separated on wideband (0–20 kHz) but not narrowband (0–8 kHz) data. In Fig. 5b, spectral characteristics are uniform across animals in wideband data but vary significantly between animals in the narrowband data. These characteristics make wideband acoustic recordings necessary for our approach to the automated detection and classification of bite events.

3.2. Acoustic contamination of recordings

3.2.1. Cross-contamination from other bite sounds

A total of 126 field observations were made of animal proximities while grazing. 7.1% of these observations involved two or more animals closer than 1 m. This yields an estimated interaction rate of 3.5% for two or more animals closer than 0.5 m.

3.2.2. Contamination from non-target sounds

A non-target noise event can intrude in two ways: (1) as a spurious event mistaken for a target event by meeting the acoustic detection criteria and (2) as a contaminating event overlaying a valid target event and spuriously increasing its energy level. Table 2 summarizes non-target noise sources, their capacity for spurious detection, and their magnitude of interference based on energy level and frequency of occurrence. None of these sources has sufficient energy within our detection band (17–22 kHz) to be spuriously detected as a bite event. Intermittent sources (such as animal vocalizations, aircraft, etc.) have a low rate of occurrence and will not significantly contaminate the data set. Continuous sounds such as crickets, when present, will contaminate every event in the data set. We calculated bite energy to cricket energy ratio and the resulting spurious increase in measured bite energy as a percentage error. We obtained bite energy to cricket energy ratios of 105.1 and 30.2 for our two samples. The worse of these would increase bite energy by $(1 + 1/30.2)/1 = 1.033$, for a percentage error of 3.3%. We consider this error level acceptable in our study.

3.3. Calibration and validation of automated bite detection

Automated bite detection was validated through: (1) comparison of manual and automated bite counts on multiple data files

Table 2
Summary of non-target noise sources.

Noise source	Detected as spurious event	Broadband energy relative to bite events	Frequency of occurrence relative to bite events
Crickets	No	Low	High in summer/fall
Flies	No	Low	Low in summer/fall
Cattle vocalizations	No	High	Low
Birds	No	Low	Low
Jet aircraft	No	Medium	Low
Farm equipment	No	Medium	Low
Road traffic	No	Medium	Low

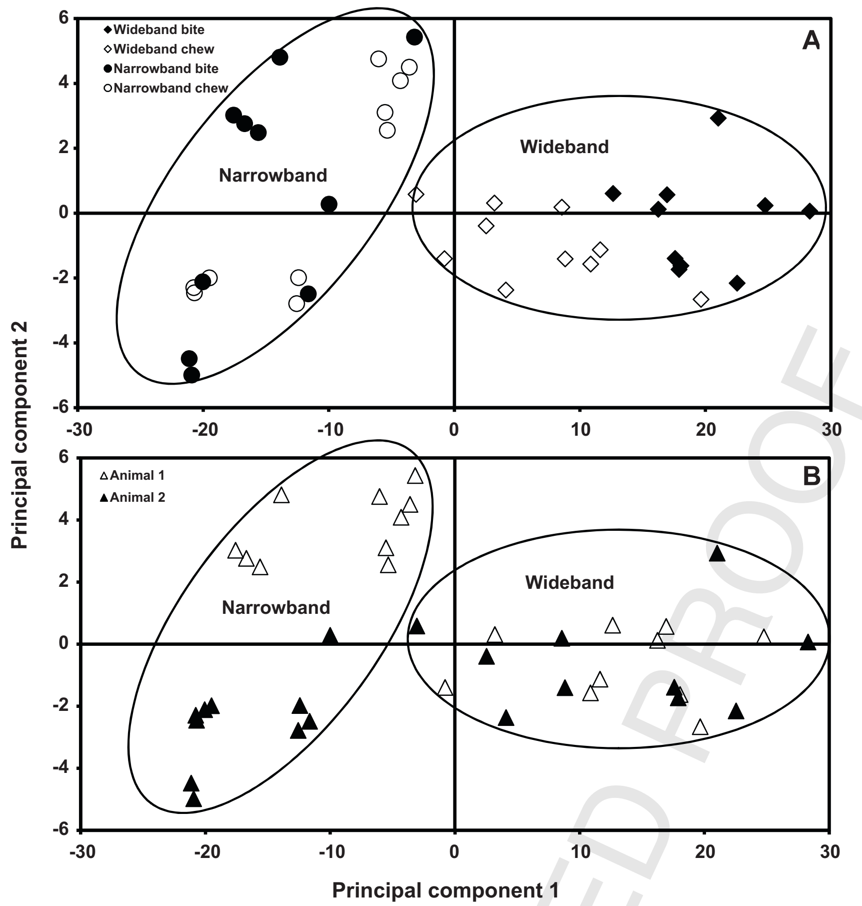

Fig. 5. Results from principal component analysis of frequency spectra comparing: (a) bites and chews from wideband and narrowband recordings and (b) the same data distinguished according to the two animals used in the analysis.

and (2) a visual event by event comparison of automated and manual bite events for one 5 min recording. Manual bite counts were validated by two independent trained observers, whose results correlated closely ($r = 0.99$; $P = 0.0001$; $n = 12$). Fig. 6 illustrates the comparison of manual and SIGNAL-derived total bite counts for three steers over 60 min of data. Although manual and automated bite counts differed by small amounts, the repeated measures analysis of variance indicated no significant difference ($p = 0.84$) between the two techniques and no significant interaction ($p = 0.53$) between count technique and individual steers. The automated bite counts exhibited a residual error of 9.1% relative to the manual bite counts. In the event by event comparison, SIGNAL identified 154 true bite events (true positives; TP), detected 2 events that were not bites (false positives; FP) and missed eight manually identified bite events (false negatives; FN). SIGNAL delivered a true positive detection rate or sensitivity (correctly detected bites/total true bites) of 0.95 (TP/(TP + FN)) and a positive predictive value (correctly detected bites/total detected bites) of 0.99 (TP/(TP + FP)) (Suojanen, 1999).

4. Discussion

Our acoustic monitoring system recorded and processed acoustic recordings of grazing activity in steers under free-ranging conditions, including identifying, classifying and quantifying ingestive events. Characteristics of the system that contribute to its successful trials include: (1) a light-weight, sturdy, halter-mounted digital recorder and microphone that had no observable impact on grazing behavior, (2) CD-quality digital recordings (44.1 kHz, 16-bit) that included the full frequency range up to 22 kHz, and (3) SIGNAL software that could utilize the high-frequency characteristics of bite sounds to automatically detect and measure bite event parameters from digital recordings of any length.

4.1. *Importance of wideband acoustic recordings*

Previous work has distinguished bites and chews based on temporal characteristics and audible differences in sound quality (Laca and WallisDeVries, 2000). Previous audio recordings of biting and chewing events relied on forehead-mounted, inward-facing microphones (Laca and WallisDeVries, 2000; Ungar and Rutter, 2006; Galli et al., 2006; Milone et al., 2009) and were apparently limited in frequency range. For example, Laca and WallisDeVries (2000), show bovine bite and chew spectra limited to approximately 6 kHz, while Milone et al. (2009) show sheep bite, chew and chew–bite spectrograms limited to approximately 8 kHz.

Our approach to automatically detecting bite events and distinguishing them from chews is founded on the fact that although the spectral profiles of bites and chews are similar below 8 kHz, they differ significantly in the 10–20 kHz range. Wideband acoustic data extending to 22 kHz (Fig. 4a) show bite and chew events as spectrally different and distinguishable, while narrowband data limited to 8 kHz show bite and chew events with similar spectral characteristics (Fig. 3b). Fig. 3a and b represents the same bite and chew events and are normalized for root-mean-square (RMS) power to emphasize differences in spectral distribution. Non-normalized plots (data not shown) depict an even greater bite–chew difference in the 10–20 kHz range, further increasing the separability of bites and chews in acoustic data. For this reason, our project is

Please cite this article in press as: Clapham, W.M., et al., Acoustic monitoring system to quantify ingestive behavior of free-grazing cattle. Comput. Electron. Agric. (2011), doi:10.1016/j.compag.2011.01.009

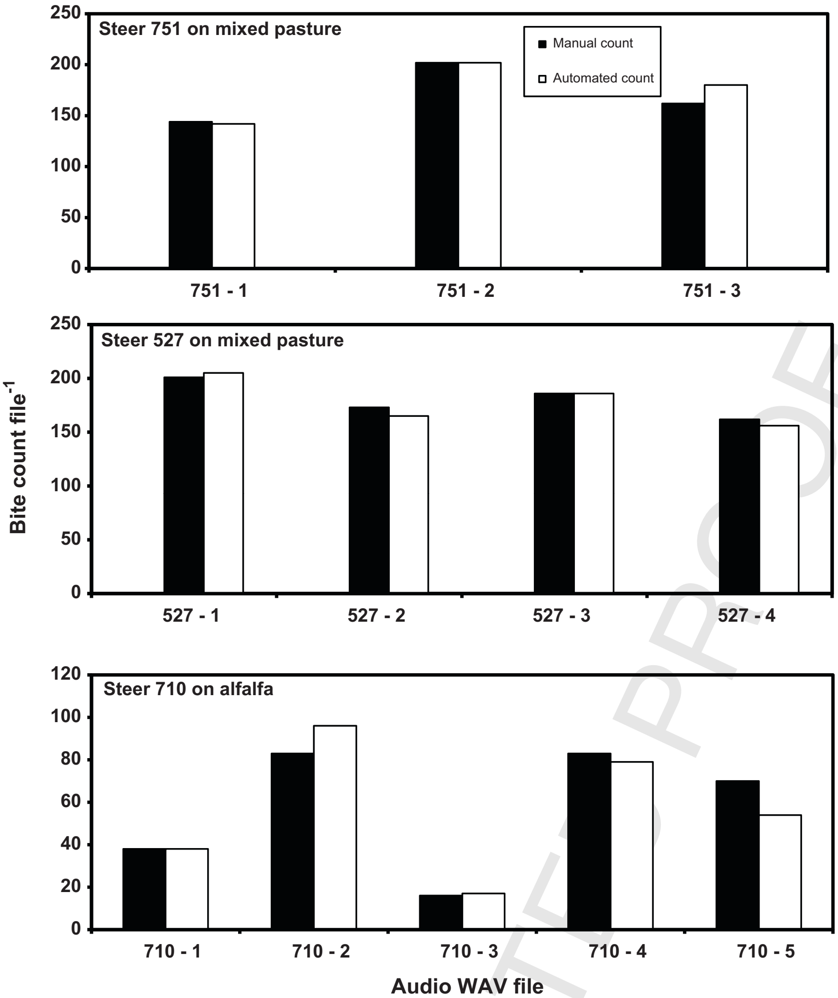

Fig. 6. Comparison of the number of bites detected by manual review of audio/visual recordings vs. automated processing of the audio portion of the same recordings. Each of the 12 segments is 3–5 min in length.

based on full frequency range acoustic recordings. The distinction between wideband and narrowband data is further confirmed by principal component analysis (Fig. 5). These data also suggest that accurate automated bite–chew differentiation would be difficult using narrowband data and possibly inconsistent among animals.

4.2. Acoustic contamination of recordings

In a manual analysis of individual events, spurious and contaminated events can be identified and excluded. However, in an automated analysis, any event meeting the mathematical selection criteria will be included in the measured data set, whether a valid, spurious or contaminated event. We therefore surveyed the range and modalities of non-target noise sources with two questions in mind: (1) can a non-target event be spuriously accepted and (2) if overlayed on a target event, what would be the quantitative impact on event energy. Our goal was to estimate the total impact of non-target acoustic events on measured bite energy.

Contamination from bite sounds of adjacent animals has been raised as an important concern (Ungar and Rutter, 2006). Our analysis indicates that even at a high stocking rate (40 animals per ha), only 3.5% of recorded bites would be contaminated at a level of 1% or greater (resulting from animals within 0.5 m of each other). However, our calculations do not account for three factors that may further reduce contamination: (1) the reluctance of these large animals to bring their heads within the 0.5 m critical distance of each other; (2) acoustic shadowing when the animals' heads are parallel but opposite in orientation. In half of these instances the recording microphone, mounted on the side of the jaw, will be acoustically shadowed from the sounds of the adjacent animal by the head of the wearer, and (3) temporal dispersion of bite events; since bites occupy less than 30% of recorded duration, energy contamination will be reduced proportionally. Considering these factors, we estimate that less than 1% of our bite events will have an energy error of 1% or greater due to contamination. (Note that contaminating bites that do not overlay a bite event are rejected by the threshold setting of the bite event detector and do not enter the data stream.) Cross-contamination levels may fluctuate with stocking density, pasture geometry and herd behavior. We expect to evaluate these assumptions further as grazing dynamics change across the season.

We analyzed contamination from non-bite noise events in two cases. First, non-bite events that do not overlay a bite event will not meet the spectral profile of a bite and will be rejected by the bite detector. Second, when non-bite noise events do overlay bite events, we estimated the resulting corruption of measured bite energy. These events divide roughly into high energy events that occur rarely (such as aircraft flyovers and cattle vocalizations) and low energy events that occur frequently or continuously (such as crickets). We calculated bite energy corruption due to a continuous non-target source, cricket sounds and found the result was a small percentage error. At the other extreme, the energy level of an aircraft flyover would invalidate any simultaneous bite events, but if a flyover occurs once per hour and shadows 10 out of 1000 recorded bites during the incident, the net corruption is again in the 1% range.

4.3. Automated bite detection

When we began the effort to automate the identification and classification of the sound data, we capitalized on the fact that in an acoustic recording of grazing, bite events had significant energy between 17 and 20 kHz, a region of the sound spectrum with little background noise in pastoral settings. This became the foundation of our acoustic bite event detector, which we programmed to focus on the 17–22 kHz range (Fig. 3). With the detection system calibrated for a given animal and forage, our data show no significant difference between bite counts derived from manual classification based on video/audio recordings and automated classification using SIGNAL. In practice, our system will require periodic manual calibration. Further tests are needed to determine the frequency of calibration, but calibration will almost certainly be required when animals are moved to a new forage resource, e.g., from mixed pasture to alfalfa. After the calibration procedure is completed, SIGNAL can process long files rapidly and with high accuracy.

4.4. General considerations

One limitation of the digital recording system is data storage capacity for the 44.1 Hz, 16-bit recordings. 32 GB SD memory cards can accommodate 48 h of continuous data recording, but the power supply must be increased to accommodate that duration, and a larger power supply increases the equipment's footprint on the livestock. A more promising approach is to implement the detection, classification and measurement algorithms on an embedded processor and store this dramatically reduced data set instead of recorded acoustic waveforms.

Development of a method to estimate grazing livestock intake is a goal that has been long sought after. Estimating forage intake is a vital step toward integrating animal performance and forage management in grazing systems and is important to measures of performance efficiency. Our recording and automated processing system solves major problems in estimating ingestive events in grazing livestock, namely, recording extended periods of free-grazing, automatically classifying bite and chew events and quantifying relative energy per bite.

Acknowledgements

We thank Nathan Wade Snyder for his assistance during all phases of this research, especially his critical roles with hardware fabrication and data management; Keith Galford and Danny Carter for their contributions to animal care and management; and Edward Pell and Shane Clarkson of the West Virginia University Willowbend Farm for their cooperation and assistance during this project. We extend appreciation to Dr. Emilio Laca for providing sound samples. Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture (USDA). The research was funded in part by the USDA-Agricultural Research Service and is part of a regional initiative, Pasture-Based Beef Systems for Appalachia, a collaboration among USDA-ARS, Virginia Tech, West Virginia University, and Clemson University.

References

Chambers, A.R.M., Hodgson, J., Milne, J.A., 1981. The development and use of equipment for the automatic recording of ingestive behaviour in sheep and cattle. Grass Forage Sci. 36, 97–105.

Champion, R.A., Rutter, S.M., Delagarde, R., 1998. Recent Developments with the IGER Behaviour Recorder. In: Proceedings of the Ninth European Intake Workshop , IGER, North Wyke, November, pp. 31–34.

Cordova, F.J., Wallace, J.D., Pieper, R.D., 1978. Forage intake by grazing livestock: a review. J. Range Man. 31, 430–438.

Galli, J.R., Cangiano, C.A., Demment, M.W., Laca, E.A., 2006. Acoustic monitoring of chewing and intake of fresh and dry forages in steers. Anim. Feed Sci. Technol. 128, 14–30.

Griffiths, W.M., Alchanatis, V., Nitzan, R., Ostrovsky, V., Ben-Moshe, E., Yonatan, R., Brener, S., Baram, H., Genizi, A., Ungar, E.D., 2006. A video and acoustic methodology to map bite placement at the patch scale. Appl. Anim. Behav. Sci. 98, 196–215.

Laca, E.A., Ungar, E.D., Seligman, N.G., Ramey, M.R., Demment, M.W., 1992. An integrated methodology for studying short-term grazing behavior of cattle. Grass Forage Sci. 47, 81–90.

Laca, E.A., Ungar, E.D., Demment, M.W., 1994. Mechanisms of handling time and intake rate of a large mammalian grazer. Appl. Anim. Behav. Sci. 39, 3–19.

Laca, E.A., WallisDeVries, M.F., 2000. Acoustic measurement of intake and grazing behaviour of cattle. Grass Forage Sci. 55, 97–104.

Liu, X., Tan, J., 1999. Acoustic wave analysis for food crispness evaluation. J. Texture Stud. 30, 397–408.

Milone, D.H., Rufiner, H.L., Galli, J.R., Laca, E.A., Cangiano, C.A., 2009. Computational method for segmentation and classification of ingestive sounds in sheep. Comput. Electron. Agric. 65, 228–237.

Rutter, S.M., Champion, R.A., Penning, P.D., 1997. An automatic system to record foraging behavior in free-ranging ruminants. Appl. Anim. Behav. Sci. 54, 185–195.

Rutter, S.M., 1998. Graze: a program to analyse recordings of the jaw movements of ruminants. In: Measuring Behavior '98, 2nd International Conference on Methods and Techniques in Behavioral Research , Groningen, The Netherlands.

Suojanen, J.N., 1999. False false positive rates. New Engl. J. Med. 341, 131.

Ungar, E.D., 1996. Ingestive behaviour. In: Hodgson, J., Illius, A.W. (Eds.), The Ecology and Management of Grazing Systems. CAB International, Wallingford, pp. 185–218.

Ungar, E.D., Rutter, S.M., 2006. Classifying cattle jaw movements: comparing IGER behavior recorder and acoustic techniques. Appl. Anim. Behav. Sci. 98, 11–27.

Zdunek, A., Bednarczyk, J., 2006. Effect of mannitol treatment on ultrasound emission during texture profile analysis of potato and apple tissue. J. Texture Stud. 37, 339–359.