prpn SIG

SPEAK OUT!

THE NEWSLETTER OF THE IATEFL PRONUNCIATION SPECIAL INTEREST GROUP

September 2009 Issue 41

Price £5.00
Free for PronSIG members
ISSN 1026:4345

www.iatefl.org

Grounding stress in expiratory activity

Piers Messum

This is a report of my presentation at the recent Cardiff IATEFL conference, where I argued that we must understand English stress not only in the form it takes in adult native speakers but also in its form in children. We should then use this developmental model for our teaching. Our learners will have to embody stress in a way that will not have been demanded from them before (in most cases), but will then have a chance of pronouncing English more authentically and easily. Up to now, I don't think we have given them the tools to do this.

In this article I will use 'stress' to refer to (actual) sentence stress, rather than (potential) lexical stress, and to refer specifically to 'stress accent' (Beckman 1986:1): the form of stress found in West Germanic languages, including English, where the correlates of stress include extra loudness, and not just pitch movement and increased length (these last two being the principal correlates found in the routine stress/prominence mechanisms of, for example, French and Japanese.)

Sentence stress in English

What is stress? Linguistically, it is a way of modulating attention at the level of the syllable (De Jong 2000). Phonetically, stress can be looked at from the different points of view of production and perception (yielding its correlates and its cues). Its developmental aspect has had less attention paid to it, but I will consider this later.

Within phonetics there has been a historical change of emphasis regarding stress, from production to perception. (Jensen (2004) has a good recent summary of this movement.) There has also been continuing uncertainty about the phenomenon.

Before 1950, phoneticians might have generally agreed with Jones (1918: 245): 'Stress may be described as the degree of force with which a sound or syllable is uttered. It is essentially a subjective action.'

After 1950, however, various experimental studies seemed to undermine this. Fry's work on the perception of pairs of words like *'import* and *im'port* relegated loudness to a supporting role among the cues to stress, where it had previously been seen as the most important result of the greater effort postulated. Adams and Munro (e.g. Adams 1979) could not find the expiratory muscle activity that was an expected correlate of stress. And Ohala (e.g. Ohala 1990) explained slight increases in subglottal pressure that accompany the production of routine sentence stress as being the result of back pressures rather than greater expiratory drive (the expiratory activity of the muscles of the respiratory system).

All of these conclusions have, in fact, been challenged by more recent theoretical and experimental work (see Messum 2007 for details), but their legacy has been the uncertainty I mentioned in the field. Notice, for example, the qualification that Roach (2002) signals with the word 'likely' in the following description:

> It seems likely that stressed syllables are produced with greater effort than unstressed, and that this effort is manifested in the air pressure generated in the lungs for producing the syllable and also in the articulatory movements in the vocal tract. These effects of stress produce in turn various audible results: … [pitch … length … loudness].

How stress is learnt

Let us for now put the question of what stress is to one side, and ask about how it is replicated. Generally, there are at least three ways that a learner can pick up an aspect of pronunciation from a model: by what I will call acoustic matching, acoustic pattern matching and action matching.

By 'acoustic matching' I mean something like mimicry, although the accuracy of this will always be limited by anatomical differences between people – for example in vocal tract sizes – that make exact acoustic matching impossible.

By 'acoustic pattern matching' I mean that the listener first identifies a pattern within the model's utterance, and then tries to reproduce this pattern with his or her own voice. So, for example, rather than matching the exact trajectory through which an intonation contour has moved, the listener abstracts the fact that the model speaker has moved from low to high and does the same.

Finally, by ‘action matching’ I mean that the listener recovers the gesture (motor activity) that the speaker has made (directly from the acoustic output it creates under some theories of speech perception, alternatively by inference and/or from visual cues), and then performs that same gesture.

The distinction between the first and second of these is not explicitly taken into account in most pronunciation teaching, and would be a good subject for a different article. For now, though, I would suggest that there is also a working or practical assumption in much pronunciation teaching that the connection between a given acoustic output and the gesture(s) made to produce it is one-to-one. So if the student is producing an acceptable output then we allow ourselves to assume that he must be, or is probably, doing it in the way that native speakers do.

Relating this to stress specifically, even some of the most production-oriented pronunciation textbooks make this working assumption.

For example, in one textbook that I like for many other reasons, Underhill (1994: 52) gives a 'practical definition of stress', describing it as an 'increase in lung power.' He then says that this causes three distinguishable acoustic results (extra loudness, extra length and pitch changes) and two other correlates (clearer vowels and more extensive articulator movements). Having identified these five variables, Underhill says: '[I]t follows that when you are trying to help your learners to produce a clearer word stress you can choose to work on whichever feature the learner finds easiest to control, or whichever feature seems to guide the learner towards the most "English" sound …'

This advice is actually similar to that given by those authors who omit any mention of increased expiratory activity in their treatment of stress, and who implicitly assume that a learner attempting to match the acoustic cues for stress will replicate stress successfully.

Surely it would be better if we demand of our learners that they increase their expiratory drive itself. In the next sections, I will consider how loudness is created and how children learn stress, and this will then give us three reasons for taking this direct approach:

it’s doing what adult native speakers probably do, some or all of the time;

the alternative – copying the loudness correlate – will probably NOT lead to what native speakers do;

children probably learn stress as greater expiratory activity, so we should ask our (older) learners to match this action itself rather than trying to match the consequences.

Even if one is unconvinced about the role of the respiratory system in routine sentence stress in adults, then the third reason will, I hope, be sufficiently convincing in itself.

How loudness works

Gauffin & Sundberg (1989) describe how it is the abruptness with which the flow of air is cut-off as it passes through the vocal folds that determines the loudness of the voice signal. The well-defined pressure pulse that is created excites the rest of the vocal tract very effectively (figure 1).

The variable that is measured to indicate this, the Maximum Flow Declination Rate (MFDR), can be increased in a variety of ways (Stathopoulos & Sapienza 1993). One is for expiratory drive to be increased and the tension of the vocal folds ‘tuned’ in accordance with this. Another, however, is for a speaker not to change his expiratory drive, but just to tense his vocal folds more.

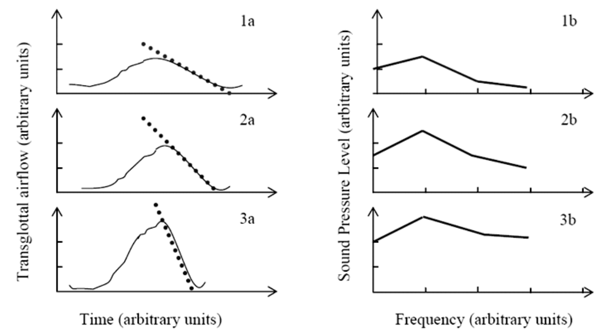

Figure 1. From Gordeeva (2005). Original caption: 'Variations in the flow glottogram of a single cycle (left part of the diagram) when a speaker was instructed to increase phonatory loudness (conditions 1a to 3a from soft to loud). Right part of the diagram represents the acoustic consequences of such increase in the radiated spectrum (2nd and 3rd ticks on the horizontal axes show frequencies between 2 and 3 kHz) (adapted from Gauffin and Sundberg, 1989).'

For most languages, speakers do not need to make regular transient increases in loudness and hence (perhaps) transient changes in expiratory drive. For a student with such a linguistic background, a purely laryngeal adjustment may be the most convenient way to acoustically match a model English speaker. Actually doing what native speakers may do – making quick increases in expiratory drive – would be an unfamiliar way of using his speech breathing apparatus, and probably difficult and uncomfortable. So this student may satisfy himself and his teacher in class that he can do a 'stress' exercise, but the manner in which he does it is probably not transferable to unmonitored speech.

How children replicate stress

There have been few studies of the development of stress in children, but Kehoe et al (1995) looked at how the acoustic correlates develop, and found that subjects in the age range 1;6 to 2;6 produce changes in pitch, duration and amplitude. In particular, they found that, 'All subjects marked the stressed syllable with higher overall intensity than the unstressed syllable. There was a tendency for the intensity difference to increase with age.'

To do this, children must be increasing their MFDR. But Stathopoulos (1995: 78) suggests that the option of a purely laryngeal adjustment to achieve this may not be available to them, because of developmental differences in their physiology: 'One noticeable difference between children and adults is that [children] do not use their laryngeal mechanism in the same way to increase vocal intensity… Maximum flow declination rate, for children, may be controlled by the respiratory system alone – through increases in subglottal pressure.'

This general understanding, that we need to see speech development in children within their particular anatomical and physiological context, is also made by Moore (2004: 193): 'Many of the characteristics, properties, and capacities incorporated into the adult model [of speech] are simply not present in, nor attainable by, young children… [T]he vocal folds appear to lack the essential biomechanical properties that would permit modulation of… intensity using adult-like strategies.'

So to increase the loudness of his speech, a child probably has to increase his expiratory drive. And Moore goes on to make another relevant point: 'More significantly, these differences reinforce the notion that speech is developed using a mechanism that is fundamentally different from that of the mature, target system.'

Another part of the mechanism that differs from the adult system is a child's speech breathing. When an adult inhales, he does so against the resistance that the tissue of his chest wall offers. This tissue is like the rubber skin of a balloon. When it has been stretched and inspiratory muscular action ceases, its elasticity naturally generates pressure inside the lungs. In fact this 'recoil' pressure provides most of the pressure that an adult speaks on.

A child's chest wall is much more compliant ('floppy') than an adult's. So inhalation gives the child two lungs full of air, but not air that is passively pressurised to a level he can use for speaking (see figure 2.) Another difference is that a child's vocal folds need greater driving pressures than adult ones, so the child must actively create 10, 12 or more centimetres of water (cmH_2O) of pressure to speak on.

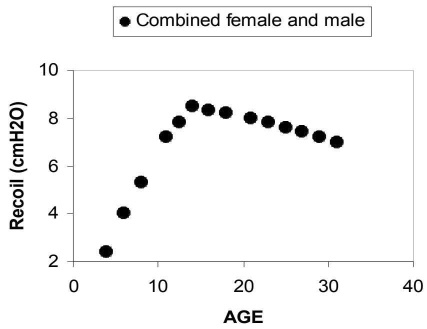

Figure 2. Redrawn from Stathopoulos (2000). Adults typically speak on subglottal pressures of 6-8 cmH_2O, similar to the recoil pressure they naturally generate at the end of an inspiration. Children speak on higher pressures, but naturally generate minimal recoil pressure. They must supplement this with voluntary expiratory muscle activity.

If we combine these two ideas – the need for expiratory activity for any increase in loudness, even transient increases, and the need to actively create subglottal pressure – then a model of child speech breathing for English emerges which is very different from the conventional adult model. Instead of a smooth delivery of power, a child will have to use pulsatile expiratory activity to speak English.

The underlying nature of stress

Understanding this helps to explain why stress production in adults has been so difficult to pin down. It may be that some or all adults can make routine stress in English with minimal or no increase in expiratory drive. However, the echoes of the childhood strategy will remain, and adult speakers will vary to the extent that they still use greater expiratory activity for stress and in the circumstances when they do this.

From a phonetic point of view, this seems to be quite significant. Understanding that stress develops in children as something very different from stress in its adult form may resolve some longstanding problems with stress.

Pedagogically, it suggests that we may have more success than at present by making sure that older learners replicate stress not through attempts at acoustic matching or acoustic pattern matching, but by action matching; with the action in question being the muscular activity that leads to greater expiratory drive.

The Accent Method

One way to achieve this might be to adapt exercises devised for the Accent Method (AM), a therapeutic programme used in Scandinavia, Germany, the UK and elsewhere by speech and language therapists.

AM exercises mainly aim to improve the 'support' that the respiratory system provides for sound production, and for the production of 'accented' (i.e. stressed) syllables in particular. However, the approach takes a holistic view of voice, so its use is not restricted to clients who have problems with their speech breathing as such. In fact, quite the opposite. The AM seems to be typically used with problems like stuttering, which on the face of it is a disorder related to articulation. AM practitioners would explain their success by pointing to the close functional, anatomical and developmental links between, for example, the diaphragm and the larynx. By retraining a client's speech breathing to function in a healthy way, various problems which appear to be located elsewhere resolve themselves.

The most authoritative account of the AM is found in Thyme-Frøkjær & Frøkjær-Jensen (2001), with shorter accounts in Kotby (1995) and Dalhoff & Kitzing (1987). I would hope that at least the following aspects of the programme will be transferable to language teaching:

- Exercises which sensitise students to how their speech breathing functions (see Fig 3).
- Exercises which encourage the use of the abdominal musculature to support stress (creating transient changes in loudness).
- Sample texts which provide practice in this way of speaking.

In the UK (and perhaps elsewhere) AM exercises have already been adapted for use in training professional voice users like singers, particularly as a result of interdisciplinary collaboration within the British Voice Association. A good account of this can be found in Chapman & Morris (2006).

In Cardiff I demonstrated some AM exercises. In the main ones, the instructor makes various sounds which start soft, increase in loudness and then reduce again. When following these, the advice given to students is that, '… the increase of intensity should not be realised by tensing the laryngeal muscles, but by an accentuated acceleration of the exhalatory air push, caused by a slight contraction of the abdominal muscles…The aim of the accent exercises is… a momentary acceleration of the exhaled air flow.' (Dalhoff & Kitzing, 1987)

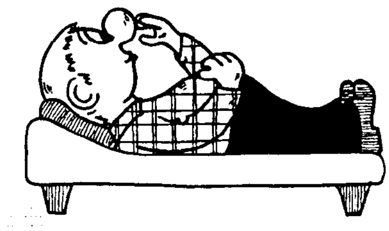

Figure 3. Lying on one's back forces a style of breathing where the abdomen rises and falls, and makes this very apparent. Illustration from Thyme-Frøkjær & Frøkjær-Jensen (2001).

So although students hear a sound that gets louder and softer, they are encouraged not to copy this as such, but to infer that the instructor is using his abdominal muscles in the way described and to match this behaviour. The change in loudness comes naturally as a consequence, rather than being the target. Hence old habits of overuse of the larynx to create changes of loudness are held at bay.

Conclusions

There will be other ways that our students can work on expiratory activity directly, and this article is partly an invitation for collaboration and exploration. I hope to report at the IATEFL conference next year on my own experiences, but would welcome contact with anyone else interested in this area in the meantime.

The goal is for our students to have a way of speaking English (1) that is healthy and authentic, (2) that is grounded in physical activity that can be automatised and successfully transferred from the classroom to free speech, and (3) that may enable the emergence of timing phenomena such as English 'rhythm' without the need for explicit instruction (Messum 2008a, 2008b).

Piers Messum is a freelance teacher from London who has also taught in France and Japan. His articles (and some further notes on this article) are available at:
http://p.messum.googlepages.com/home.
He can be contacted at p.messum@gmail.com

References

Adams, C. (1979) *English Speech Rhythm and the Foreign Learner*. The Hague: Mouton.

Beckman, M.E. (1986) *Stress and Non–Stress Accent*. Dordrecht: Foris (Netherlands Phonetics Arch. No. 7).

Chapman, J. & Morris, R. (2006) Breathing and support. In Chapman, J. (Ed.) *Singing and Teaching Singing.* San Diego: Plural.

Dalhoff, K. & Kitzing, P. (1987) Voice therapy according to Smith. Comments on the Accent Method to treat disorders of voice and speech. *Journal of Research in Singing* 11, 15.

de Jong, K.J. (2000) *Linguistic stress and attentional modulation*. Presentation at Cognitive Science Colloquium, University of Rochester.

Gauffin, J. & Sundberg, J. (1989) Spectral correlates of glottal voice source waveform characteristics. *JSHR* 32, 556–565.

Gordeeva, O.B. (2005) *Language Interaction in the Bilingual Acquisition of Sound Structure*. PhD thesis, Queen Margaret University College.

Jones, D. (1918) *An Outline of English Phonetics*. Cambridge: Heffer.

Kehoe, M., Stoel-Gammon, C. & Buder, E.H. (1995) Acoustic correlates of stress in young children's speech. *JSHR* 38, 338–350.

Kotby, M.N. (1995) *The Accent Method of Voice Therapy*. San Diego: Singular.

Messum, P.R. (1998) The Accent Method. *Speak Out!* 22, 30–33.

Messum, P.R. (2007) *The Role of Imitation in Learning to Pronounce*. PhD thesis, London University.

Messum, P.R. (2008a) Embodiment, not imitation, leads to the replication of timing phenomena. *Acoustics 08.* Paris: SFA/ASA/EAA. 2405–2410.

Messum, P.R. (2008b) If children don't learn to pronounce by imitation, how should we teach older learners? *Speak Out!* 39, 16–21.

Moore, C.A. (2004) Physiologic development of speech production. In B. Maassen, R. Kent, H.F.M. Peters, P. van Lieshout & W. Hulstijn (Eds.) *Speech Motor Control.* Oxford: University Press.

Ohala, J. (1990) Respiratory activity in speech. In Hardcastle, W.J. & A. Marchal (Eds.) *Speech Production and Speech Modelling.* London: Kluwer.

Stathopoulos, E.T. & Sapienza, C. (1993) Respiratory and laryngeal function of women and men during vocal intensity variation. *JSHR* 36, 64–75.

Stathopoulos, E.T. (1995) Variability revisited: an acoustic, aerodynamic and respiratory kinematic comparison of children and adults during speech. *J. Phon.* 23, 67–80.

Thyme-Frøkjær, K. & Frøkjær-Jensen, B. (2001) *The Accent Method*. Bicester: Speechmark.

Underhill, A. (1994) *Sound Foundations*. Oxford: Heinemann ELT.