

# How not to fix a problem: misapplications of pronunciation theory

Sam Bamkin

*De Montfort University, UK*

Prevalent pronunciation teaching methodology is traced through coursebooks to teacher handbooks, and directly to academic theory, unmediated by theories of pronunciation teaching or learning. The missing aspects of pronunciation teaching are then sought, requiring a complete realignment of the theory in line with other aspects of language teaching. The theoretical implications of this paper expose a possible mass confusion between the teaching of receptive and productive skills. Premature application of theory is traced through various patches and modifications not amounting to fixing the fundamental problems. These various modifications together map an explanation of the confusion rooted in the compounding of earlier misapplication. Following the exposed discrepancies, a more rational model, also in line with other areas of language learning, is proposed. Finally, some possible requirements of future directions for the practical implementation of these are briefly outlined.

## Introduction

Pronunciation is becoming more and more widely discussed in the British English Language Teaching (ELT) training community. The British Council and university training centres that follow suite are repeating the motto: *focus on pronunciation*. Although consensus exists on its importance for speakers of other first languages (L1), how a teacher *focuses on pronunciation* remains an unresolved and minimally explored question. More worryingly, widespread misdirection in the application of theory prevails in many quarters of the profession. Whilst the foundations remain shaky, any development of teaching material granting student, and teacher, satisfaction remains unlikely. The current paper examines whether the very roots of pronunciation teaching could cause the problems inherent in many current approaches.

Despite peaks of emphasis, most notably during the audio-lingual influenced period, before the current communicatively induced lull (Leather, 1983; Levis, 2005; Setter and Jenkins, 2005; Derwing and Munro, 2009), little development has occurred in pronunciation teaching or learning research in recent years. This is overshadowed by leaps and bounds in the descriptive frameworks of pronunciation and scholarly interest in the value of pronunciation. The disparity is summarised most cogently by Derwing and Munro:

An extensive, growing literature on L2 [second language] speech has been published in journals that focus on speech production and perception [...] Yet this work is rarely cited or interpreted in teacher-oriented publications (2005, p. 382).

This dearth of relevant research persists today (Leather, 1983; Terrell, 1991; Elliott, 1995, 1997; Derwing *et al.*, 1998; Levis, 1999; Lord, 2005). The few islands of recent

insightful examination here discussed remain the exceptions in an ocean awash with blind application of intuitive guidance (Levis, 2005).

After defining the parameters of analysis, the general rubric of current and prevalent pronunciation teaching methodology is expounded, prompting an analysis of the extent to which its grounding in descriptive pronunciation is mediated by theory of production, reception, and their learning. A survey of commonly utilised classroom materials is then presented to illustrate these prevalent trends. Coupled with ambiguous experimental results and recent descriptions of pronunciation learning, a more considered approach is proposed. The basic tenets are to treat the learning process as holistic: setting targets, or showing *what* to learn, next instructing from explicitly to inductively, or showing *how* to learn, and finally finding ways to practice communicatively to enable sustained *rehearsal* to naturalise the forms.

## Pronunciation theory and learning

Pronunciation is the evanescent manifestation of acoustic language, or the ability to produce or receive it accurately. Since the mechanisms of the ear are static, the mind is responsible for receiving the acoustic manifestation for processing. *Receive* is here used in the specific sense of recognising linguistics forms or words whilst they are heard. The *knowledge* part of pronunciation, no doubt a psychological reality, appears to be shared by both the productive and receptive faculties, corresponding to audio, muscular, and indeed orthographic manifestations only as an abstraction.<sup>1</sup>

The term *form* is utilised to describe a language feature or isolated component of usage. The isolatable nature reflects *formal* linguistics, which comprises of descriptive literature on each linguistic element, or form. The current focus is upon learning productive skills, although its close and multifarious relationship with receptive ability, along both theoretical and pedagogic dimensions, requires extensive discussion of both. Pedagogic pronunciation is here taken to include all aspects of pronunciation, although these are often narrowed in scope depending on the course content decisions within the educational environment in question. It has become dogmatic to place more emphasis upon supra-segmental forms – those that span entire sentences, since they are seen to carry greater bearing upon understandability. Again, this dogma has not been based upon empirical findings (Derwing *et al.*, 1998; Hahn, 2004) and only in recent years have Derwing and other researchers qualified these intuitive claims with relatively weak support with native-speaking listeners (2003; as analysed by Levis, 2005).

Two reasons support taking a broader, holistic view of pronunciation. Firstly, Flege (1981) appears to suggest that even if segmentals contribute slightly less to understanding, understanding their learning nonetheless indicates learning and development patterns. In other studies, Flege could be suspected of utilising consonant segmentals positively due to the relative objectivity that they lend to any qualitative testing required. Secondly, ‘the listener’ in many generalisations cannot be assumed to refer to ‘any listener’. The need to speak with listeners of multiple first languages, however, is growing. The question of which pronunciation aspects are salient when communicating with native-speakers or non-native speakers of third L1s remains speculation or even assumption (*cf.* for example, Jenkins, 2000; Levis, 2005; Riney *et al.*, 2005). Despite this, subscription to any assessment of importance is compatible with an approach aimed broadly at all aspects. Modern opinion on the goals of pronunciation teaching does point towards placing intelligibility as the goal of

instruction (Leather, 1983; Pennington & Richards, 1986; Morley, 1991; Hewings, 2004; Derwing & Munro, 2005). Although a contentious, and unquestionably important, question, the matter of which level learners ‘should’ be required to achieve is here avoided. Approaches to teaching pronunciation in its generality are discussed. Although examples tend to be drawn from the more extensively studied realm of segmentals, particular care is taken to extend all findings, in theory, to supra-segmentals, at least in the absence of any evidence to the contrary. Indeed, one argument below against the prevalent model is its uncomfortable accommodation of supra-segmentals. To emphasise this point, examples chosen here are, to the contrary, decisive toward including supra-segmentals.

### Analysis and survey of the prevalent model

The current and prevalent theory of tackling pronunciation is outlined and rationalised by Morley (1991). Taking a broad approach without subscription to specific methods or classroom approaches, it translates the general language learning objectives of *imitation*, *rehearsal* and *extemporaneity* to pronunciation. Although stated only that these three progress to greater desirability, they implicitly represent stages in learner development. Partially justified by a supposition that Morley discusses purely the spoken part of pronunciation and to align this progression into the general language teaching paradigm, *noticing* is demanding as a precursor to these current objectives (Pennington & Richards, 1986). The prevalent trends in pronunciation teaching do correspond to Figure 1:

Development of the skill of pronunciation			
1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>	4 <sup>th</sup>
(a)Noticing	> (b)Imitation	> (c)Rehearsal	> (d)Extemporaneity
<i>Horizontal arrows indicate stages in development</i>			

Figure 1. Development of the skill of pronunciation

Morley’s model further reflects Kihlstrom’s influential consolidation of the mapping of the neurological learning process (1984, simplified in Schmidt, 1990) in which *noticing* by attention to the sensory registers moves information to the short-term memory<sup>2</sup> in which conscious processing is possible, requiring further *rehearsal* before being memorable, later accessed without concentration or *extemporaneously*. The slight discrepancy between the two models is tentatively resolved below. For now, it should be noted that *imitation* is an extra stage required only for the pronunciation aspect of language learning. During the development of other linguistic skills, *noticing* implies the ability to imitate the essential features of the form. In pronunciation, this is not seen to be the case. Since this terminology is maintained as a useful descriptor of stages in language, whilst their content may require bifurcation and narrowing below, they are italicised.

It may be helpful to view past trends in pronunciation training as developing along the possibly complementary axes of receptive and productive skills. The simplicity of this method was encapsulated most starkly in influential texts such as Gimson's course (1975). The basis of Gimson is that once the student is directed to a form, she will then be able to isolate it. Subsequently to this, she will be able to produce the form. To utilise Morley's terms, this will become *extemporaneous* through continued *rehearsal*. Linguistic forms are presented in relatively simple, easily followed sections, with instructions to *identify*, or to *listen and repeat* various aspects, one or the other appearing five or ten times per page (a CD is provided for listening). In terms of current usage, *listen* may be seen to approximate a crude form of aiding a learner to *notice*, whilst *repeat* would approximate an aid to *imitate* and *rehearse*. Other implicit links between the terminology used for activity types, and that used for methodological foundations may be visible. However, since it is argued that the methodology lacks a foundation in the theory, the concepts are kept separate until a more complete analysis is conducted. It is suggested below that this relationship must be examined before links can be made with any confidence.

The *listen* and *repeat* components have since progressed independently along similar, and potentially complementary, paths. *Listening* appears to have assimilated the *identify* part in terms of presentation, although teacher goals generally focus on either raising awareness of a form or identifying a form of which the learner has been previously aware. How this affects the relationship to Morley's model, so aptly reflected in Gimson, is unclear. Perhaps both would be viewed as *noticing*, at least as far as it is a stage towards production since both utilise receptive skills. Browsing through the majority of historic and modern coursebooks for learners, the tasks can be seen to develop receptive skills through explanation and listening activities, then providing situations for practicing, or rehearsing the forms. Kenworthy's valuable handbook advises that:

[t]he suggestions given for classroom activities will be presented under two headings: first, activities for building awareness and concern for pronunciation in the learner; second, ways to extend and consolidate their developing skills (1987, p. vii).

A survey of teacher handbooks for pronunciation and pronunciation activity books examines the focus of activities presented for classroom usage to qualify any generalisations or assertions made regarding published methodology. The publications are selected to present materials from various viewpoints, particularly those that have gained wide recognition or are popularly found in teachers' inventories. The survey begins with Kenworthy's (1987) *Teaching English Pronunciation*, appropriately one of the latter aimed at teachers mentioned by Morley (1991). In addition, Kelly's volume in Scott Thornbury's *How to* series, *How to Teaching Pronunciation* (2000), and Hewings' volume in the *Cambridge Handbooks* series, *Pronunciation Practice Activities* (2004), are surveyed as well as Hancock's highly praised *Pronunciation Games* (1995).

Needless to say, for the most part, Kenworthy maintains her division between the previously mentioned headings, although several activities do cross the chapter boundary or develop multiple skills. Her first activities chapter contains six knowledge based activities, and 15 purely receptive based. None of these attempted to describe how to make a sound or sound pattern. The following practice chapter provides six more knowledge based activities with six listening and 11 practicing. Additionally, there are seven to train a learner to make sounds. Many are discussed

below. To classify activities, a common sense approach was taken. Some judgement was necessary to categorise activities when activities *could* be construed or easily imagined to develop along lines not explicated in the text. Only explicated procedures contributed, however. For example slur links showing that English exhibits a slight preference towards open syllables (Kenworthy, 1987; Hewings, 2004) were seen as knowledge based whereas elongating the s in '*span*' to eliminate vowel epenthesis, as in '*espan*', explicated sound making development by forming the position before expirating (Hewings, 2004). In the latter, the learner is instructed to do something specific to improve the production of a sound pattern – or to override another as the case is here. Along similar lines, a division was identified between assigning sounds to the 'ough' counting as *orthographic*, and assigning sounds to the '-ed past tense' counting as *knowledge based*, since the latter applied to a role rather than a written convention. Lip reading was counted as *contextual listening*. Although the judgements may cause delicate divergences if scrutinised, the results remain clear enough to be of value. The differentiation between *listening* out of context and in context relied partially on the stated goals or position in chapter, informed by whether the activity was framed as raising initial awareness of a form or as receiving a form with previous familiarity.

It should additionally be noted here that setting the boundaries between activities presented a challenge. 'Activity extensions' were classified as separate activities in Kenworthy (1987) because the focus often shifted, although not always. In Laroy (1995) 'stages' were separated since the focus only changed here. Further, to tilt any possible imbalance towards conservatism, any doubtful boundaries were seen not to divide activities, unless it was classified as *making sounds*: in which case, each individual aspect was classified as a whole activity. The effects of this conservatism can be illustrated by pointing out that of the 67 pages surveyed, the latter seven activities classified as *making sounds* spanned less than six pages (1987). Likewise, 17 of Kelly's 18 appear across two out of seven chapters.

All but one of Kenworthy's (1987) *making sounds* activities were directed towards imitating individual phonemes – the shortest *phonological* unit. One surprising feature of Kenworthy is the lack of facial diagrams. Those sounds that are tackled use descriptions without technical language (Kelly, 2000). These are in fact simplified models, with the potential to be backed up with facial diagrams, but entirely self-contained. The more creative procedures are discussed below. The remaining activity for aspiration, also found in Hewings (2004), uses instructions to breathe out for a short period whilst saying plosives using a sheet of paper to provide a concrete check of the airflow. Again, the technique aims to aid learners' production of various syllable-initial plosives in non-cluster positions, to be practiced later.

Kelly (2000) and Hewings (2004) both exhibit front facial diagrams to illustrate the lip positions for producing the vowels, in addition to more convention profile facial diagrams. Besides Hewings' vowel epenthesis elimination activity, one other is added to the short list of those not teaching individual phonemes activities. Kelly clearly explicates the aspects of syllable and time for diphthongs. This should be a common teaching aspect given that these outline their two defining features in general and in English respectively. Although three activities span multiple phonemes, they are clearly within the realm of segmentals. This is surprising since common wisdom suggests that suprasegmentals hold more sway. Even if native listeners alone receive more via supra-segmental cues, more instruction should have been put into teaching *how* to make these sound patterns. The final example in Kelly that does focus on supra-segmentals is discussed below.

Table 1. Categorisation of activities in four widely used teacher handbooks and activity books

Activity Category	Kenworthy 1987	Hancock 1995	Kelly 2000	Hewings 2004	Total
Knowledge based	12	23	22	26	83
Listening (total)	21 (40%)	8 (22%)	22 (23%)	29 (38%)	80 (31%)
└De-contextual listening	7	4	12	9	32
└Contextual listening	14	4	10	20	48
Making sounds (total)	7	0	18	3	28
└Individual phonemes	6 (12%)	0 (0)	16 (17%)	1 (1%)	23 (9%)
└Other sounds patterns	1 (2%)	0 (0)	2 (2%)	2 (3%)	5 (2%)
Practicing <sup>3</sup>	12 (23%)	8 (22%)	20 (21%)	24 (31%)	64 (25%)
Orthographic	1	0	15	4	20
Learning strategies and affective goals	2	0	1	2	5
Total <sup>4</sup>	52	37	94	77	260

The *knowledge based*, *listening* and *practicing sound patterns* appear in pronunciation coursebooks such as *English Pronunciation in Use* (Hancock, 2003; Marks, 2007) and *New Headway Pronunciation* (Bowler & Cunningham, 1999; Cunningham & Bowler, 2000). Those working on individual segments, facial diagrams, lip position diagrams and technical and simplified explanations, have all also found their ways into coursebook. In the interests of variety, each of these appears on virtually every page of the book series. But these are the only *making sounds* activities found. Furthermore, each coursebook reflects the general state of published material as summarised in Table 1: *knowledge based*, *listening* and *practicing sound patterns* activities are found on virtually every page, with *making individual phonemes sound patterns* confined to their designated areas (Bowler and Cunningham, 1999, the first two pages of each unit; Cunningham and Bowler, 2000, the first two pages of each unit; Hancock, 2003, Section A out of A-C; Marks, 2007, Section A out of A-D). The same patterns are found within single activities in general coursebooks. The activities are either *knowledge based* or jump starkly from *listening* to *practicing*. Again, outside of these areas, *making sound pattern* activities

are not found. These correspond to the teacher handbooks and pronunciation activity books.

With activities sensitively developed to aid the learners to become quite adept at *noticing* the forms, the prevalent model thus appears to be very effective at teaching *what* to learn for production. This would presumably enable the learner to imitate, besides in the realm of segmentals, where training *is* seen to be required to imitate. Although lying outside of the current study, some genuine changes in general ELT approaches have impacted pronunciation activities. Although not of a substantive nature in this sense, they are discussed as valuable to guide or limit any approach developed.

## **Learner Centring of Approaches**

The move toward centring any approach taken on the learner has become a recurrent rubric of texts on pronunciation. Shifting from exercises to an activity-based approach itself can be seen to attempt, with potential success, integration with a communicative model. Highlighting a communicated message through the pronunciation is frequent. Indeed, the success of Hancock (1995) may be its making standard exercises cooperative and therefore more enjoyable. Communicative credentials are arguable but the focus is undoubtedly on form. Although an important supplement to augment any methodology, this is not seen to impact the methodology proper, which is charted in the present survey.

Any approaches developed for teaching pronunciation must take a learner-centred view. In discussing the move towards learner centring, there are two points to be made. The high sensitivity brought about by the inter-relationship between pronunciation and personal identity (Laroy, 1995; Hewings, 2004; Setter and Jenkins, 2005) is worth noting, as it heightens the impact of any trends discussed. These feelings must be taken into account when developing syllabus structure or teaching materials, especially since it is becoming increasingly documented that learners are highly concerned with pronunciation. Arteaga (2000) and Yule and MacDonald (1995), for example, list many authorities opinions emphasising this trend. These opinions have since been backed up with empirical figures by Timmis (2002) in an EFL setting and Derwing (2003; cited in Derwing and Munro 2009) in an ESL setting.

Quite apart from student desires to learn, a view exists that teaching pronunciation is simply *difficult*, or, worse, ineffective. Kenworthy has termed this 'benign neglect' (1987, p. 85). This may also encapsulate genuine views that teaching pronunciation is ineffective (Chela-Flores, 2001) or a lack of confidence in teachers (Derwing *et al.*, 1998). The effects of the belief amongst teachers that pronunciation teaching is simply difficult, however, are rather more disconcerting. Summaries of teacher views highlight the widespread belief that teaching pronunciation is difficult (Kenworthy, 1987; Elliott, 1995, 1997; Setter & Jenkins, 2005), whereas calls for research often narrow this down to specific problem areas: for example, Levis on intonation (1999; Kelly, 2000), Laroy on rhythm (1995), Levis in fact describing all supra-segmentals 'an issue' (2005, p. 370). Whether genuine or not, a recognition of the difficulty of teaching pronunciation cannot outweigh the importance as an aspect of English.

The limitations on pronunciation teaching certainly appear to stem from objective judgements of who should learn, its difficulty, and possibly the level to which it should be studied. Derwing and Munro (2009, p.485) observe that:

In these days of learner-centred curricula, it seems ironic that some authorities advocate the opposite of what students want.

Any adjustment or transformation in teaching approaching, however, must continue the trends centring the approach on the learner.

### **Implications of the Prevalent Model**

Although Gimson's (1975) model can be updated to characterise the *rehearsal* stage as increasing in communicative placement, the foci remain in a similar state to those in Gimson's day: *listen* then, after some understanding, use (*practice, repeat*) for communicative purposes. Development has occurred only in the presentation of fundamentally similar methodology. Indeed, it is concluded that even Gimson's foundational pronunciation training was not built on a theory of learning, having defaulted to presenting linguistic forms followed by an expectation to produce linguistic forms. This represents that which one internet writer termed *linguistics applied*, as opposed to applied linguistics (Cauldwell, 2002). The observed half-century of stagnation in methodology, though waxing and waning in perceived importance, continues not least because Morley's (1991) model, based on Gimson's (1975), can be *construed* into Kihlstrom's (1984) accepted model alongside the training in non-pronunciation linguistic forms, although the role of *imitation*, currently only half applied with adjunct status, does not sit comfortably.

The current, prevalent methodology appears to rely on a combination of learning knowledge and *noticing*. This appears to progress naturally onto *imitation*. The proposition that *noticing* begets *imitation*, however, appears tantamount to proposing that *noticing is imitation*. This would, at least, appear logical and mirror grammatical form learning, but the inclusion of *making sounds* for individual phonemes affirms that *imitation* is seen to be different. The division could be affirmed if instruction is required specifically for *imitation* after *noticing* has reached sufficient development. There is some evidence towards this view for segmentals. Catford and Pisoni (1970), for example, conducted a remarkably controlled experiment in which participants learnt exotic sounds – those with a close counterpart in the L1 – using instruction on *noticing* alone and with instruction on *how to imitate*. The group with *how instruction* produced the new sounds significantly more accurately than the *noticing* group. The researchers were unequivocal in affirming the:

inadequacy of any purely auditory tape-recorded pronunciation-training program which relies entirely on mimicry of vowels without supplying explicit information at least on lip positions (p. 481).

In recent years, a number of researcher teachers have begun to experiment in the classroom with teaching methodology over entire language courses. Elliott (1995, 1997) and Lord (2005) have both provided empirical backing to propose that teaching pronunciation, including *how* to produce sounds, yields significantly better results than without. Although their rigorous control and results give numerous insights into the current proposal, it must be borne in mind that these extend only into the previously implementing realm of individual segmentals. Additionally these are on Anglophone learning of Spanish as a foreign language. Elliott's first experiment (1995) studied three environmental allophones. One group was given appealing class presentation including listening (*noticing*) and usage information

(*knowledge*), taught concrete articulatory rules, deductively and inductively learnt to articulate, were given immediate feedback (*making sounds*), and underwent drills (*rehearsal*). They improved significantly compared to the control group. For most of the 19 allophones, in most circumstances, Elliott's 1997 study corroborates these results. Unfortunately, reducing their parity with Catford and Pisoni, this does fail to assess the significance of the *how* instruction as divorced from the *listen* and *repeat* stages since the control group received no pronunciation training. This is mitigated slightly by a clear emphasis on *how instruction* in the report. Lord's study (2005) was largely the same as those of Elliott, different in methodology only by reducing focus on immediate feedback. One interesting factor was his inclusion of disambiguating word linking strategies: whereas any English intervocalic word boundary assimilation tends toward glottalisation or glide epenthesis to avoid hiatus, Spanish tends towards diphthongising. This type enters the realm of bi-segmental features, but again does not extend to supra-segmentals.

Derwing, Munro and Wiebe (1998) compare the instruction of *making sounds*, leaving all other aspects intact in a control group receiving no specific instruction. This was, and possibly remains, the only study of its kind in a classroom environment. The experimenters report that the 'no specific instruction' group was given pronunciation instruction, 'emphasising both receptive and productive skills' without details on how (p. 397). With Derwing, Munro and Wiebe's particular research focus in mind, this is read to approximate normal pronunciation instruction, providing a control group for the present study, although the data do not make any further reference to classroom methodology in this group. It is implicit that the control teacher excluded focusing on the form of sound patterns since the two experimental groups were excluded from focusing on the other's aspect. This could, however, have impacted detrimentally on any 'usual' pronunciation work, despite its apparent permissibility. Regardless, of the two experimental groups, those receiving supra-segmental instruction improved their ratings in more situations than those receiving only segmental instruction. Both improved their pronunciation ratings more than the control group with training that was 'normal' from this group teacher's point of view. These results appear promising but lacking complete congruence with the current focus.

The division between *noticing* and *imitation* in the prevalent model appears significantly less clear for supra-segmentals. This is because *making sounds* activities for *other sound patterns* are not found in the literature preventing any research to test the effectiveness of *how instruction* – Derwing, Munro and Wiebe's being exceptional in its possible tangential positive implication. In principle, however, in the case of the segmental evidence being explicable, Flege (1981) would appear to indicate that successful progress in the teaching of segmentals would, in principle, be similarly successful for other aspects of pronunciation. In this case, instruction in *making sound patterns* would be equally as essential for other segmental and supra-segmental features. This appears to be a particularly demanding question given the dissatisfaction with learning in these areas.

A fresh but intuitive look at pronunciation learning would appear to support this view. A learner who hears a dental fricative as distinct from other English consonants may not know how to produce it, even after watching somebody else demonstrate. In a less idiosyncratic vein, Mann (1986) established that Japanese learners are capable of distinguishing potentially salient features of the // and /r/ sounds despite failing to produce them distinctly. This problem is compounded for vowels since the articulation is difficult to see. The rounded open-mid back is often problematic, for

example. Finer features such as stronger, longer vowel quality and stronger, slightly sustained, aspirated consonant quality in stressed syllables as opposed to non-stressed content may be heard, and even contribute towards receiving prominence, without the learner knowing how to produce all aspects. The native-speaker association of these features with stress in the sense of tension has little meaning to non-native learners who may only know the linguistics definition of 'stress'. Moreover, considering differing voice settings or airflows, the new form may be very difficult to extrapolate from acoustic information alone (Honikman, 1964; Esling and Wong, 1983). Considering supra-segmental features, such as rhythm and intonation, learners can often detect features of prominence and pitch contour meanings without producing them. To provide one specific example, Nagamine comments that:

It was obvious, for instance, that my [learners] in the initial stage of learning experienced great difficulty in changing their pitch immediately after they produced a voiceless stop consonant (personal communication).

Although more obvious examples have been selected, *imitation* may not flow from receptive *noticing*. It may need to be learnt in its productive form. Chance might provide a marginal possibility, but frequently cognitive exertion is required to create new sounds or sound patterns. The mind must learn *how*. Teaching *how* to produce phonemes appears essential to their *imitation*.

Although it is clear that instruction, including *how instruction*, is significantly more effective than no specific instruction (Derwing, Munro and Wiebe, 1997, 1998) for producing supra-segmentals, as well as segmentals, it is unfortunate that no research has yet indicated empirically whether instruction in *how* to produce is more effective than the prevalent methodology of *noticing* and *rehearsal* alone. Exploring the theory behind the necessity for *imitation* for phonemes may provide a framework upon which to postulate its necessity. The general dissatisfaction with the results of pronunciation instruction being a given, any clues that could highlight deficiencies in the prevalent model should be sought.

If rightly assumed that the *rehearsal* of a productive form must be productive, no space is found in the prevalent model for further listening. Under the prevalent model, listening appears to constitute *noticing*. It would appear otherwise logical that further listening would *rehearse* this. However, *rehearsal* of a productive form must, intuitively, be productive. Besides leaving no space in the prevalent model for further listening, this creates a tension between receptive and productive development since *rehearsal* is not possible through listening. It again appears to assume that *noticing* leads to *rehearsal*, at least for supra-segmental forms, which appears counter-intuitive. The results of receptive training are examined afresh, particularly paying particular attention to its relationship with productive skill. Receptive skills themselves do constitute a vitally important skill for learners; indeed, as now examined, they are essential to identifying the *what* of production.

## **Towards a Rational Learning Model**

Not only are receptive skills important, they may also work towards improving productive pronunciation. Pimsleur (1963), using standard minimal pair discrimination training along with cross-language comparisons, found that learning to differentiate can significantly improve productive pronunciation, although not always. A repetition with consideration of more variables by Mueller and Nielzielski (1967)

yielded similar conclusions with less equivocation. Lane and Schnieder (1963) found students able to differentiate with echoic imitative responses, when the response is elicited immediately after an identical stimulus.

The immediacy of Lane and Schnieder's (1963) methodology is relevant, since it suggests that teacher presence is required for imitation. The early division of *listen* from *repeat* may represent a latent recognition that any production to become extemporaneous should not rely upon an external stimulation. Despite severely limiting the *rehearsal* time for the student, it is wholly likely that dependant *rehearsal* is marginally possible in this way, assuming that the learners are able to produce the form in question. Forms that are learnt without apparent concentration also raise issues. Prior *noticing*, innate imitative ability, or chance could be again be supported by teacher exhortation or receptive knowledge (Stevens, 1974). Although constituting possibly 70% of sounds if accuracy is ignored and even less of prosody and intonation (*ibid*; also Leather, 1983), discussing *rehearsal* development unfortunately leapfrogs the conundrum. These sounds *can be rehearsed*, and production reinforced if pronounceable. The problem, as suggested in Lane and Schnieder's (1963) study, is that many forms remain *unpronounceable*. Thus, *rehearsal* had not begun regardless of whether the teacher was eliciting echoic responses or not, unless the student approximates the form by chance.

The prevalent model suggests that sustained receptive training of struggling learners would carry over to their productive pronunciation. Neufield, on the bases of some suggestive preliminary results (1980), examined whether the relationship is this simple (1988). Comparing the receptive and productive abilities of three experimental groups: one Francophone, one Anglophone with high L2 proficiencies in French (XB) and one Anglophone monolingual with only elementary familiarity with French. Three receptive trials were conducted. In the first, recordings were analysed for foreign accent. The Francophones and XB performed with no significant gap. In the second, no deviation was found for XB identifying the phonology of sentences as 'bizarre', although a trivial gap arose when asked to isolate the 'bizarre' inappropriate allophone. In the third, XB and the Francophones were both able to isolate acceptable or unacceptable pseudo-words. XB did not correctly identify some as dubious. The Anglophone monolinguals operated at levels close to chance in all three tests. This would suggest that the XB receptive knowledge in global, as in the first test, and micro, as in the second two, levels closely approached that of the Francophones.

Besides providing strong evidence that the receptive phonological knowledge systems a learner can regularly reach near-native states, it bolsters the asymmetry between receptive and productive skills (Oyama, 1976; Neufield, 1988). Receptive ability does not beget productive ability, but is a necessary auditory *feedback loop* for production (Yule *et al.*, 1987). Although logically, imitation *can* occur by chance or *rehearsed* sporadically with a patient teacher, receptive ability, reflecting Morley's (1991) progression stages, constitutes only a *prerequisite* to *productive rehearsal* (Stevens, 1974; Kenworthy, 1987; Levis, 1999; Arteaga, 2000) – particularly achieving accurate production *rehearsal*:

The extent to which one affects the other cannot be underestimated; one needs to be able to hear a phonetic contrast before one can successfully produce it, for example (Setter and Jenkins, 2005, p. 6).

Setter and Jenkins' quote exemplifies only one dimension of the complex; likewise, only one feedback loop has drawn attention so far. A close reading of Leather (1983) places equal weight on the auditory feedback loop and two other sensory loops originating in the nerves of the articulatory organs. These comprise of the tactile, externally sensed contact with other organs; and the proprioceptive, influenced by the perception of absolute and relative organ positioning. Discussion of the sensory feedback loop appears largely confined to a later stage, during *rehearsal*. This is developed briefly to construct a comprehensive framework for the main discourse on imitation, but proceeds only on the assumption that a form is produced correctly.

When it is acknowledged that the *internal, productive* manifestation is fed back, it becomes clear that the knowledge necessary for production is not necessarily representative of the sound, but equally corresponds with a representation of muscle positioning for production. Auditory feedback may inform articulator positioning, but only indirectly. For production, the form to be *noticed* and learnt is the *imitation* itself, a manifestation in the position of muscles. *Receptively noticing* the acoustic manifestation does not directly develop the productive skill. The concept of *imitation* may prove to be undesirable, constituting a loaded term, presumptive of a causative relationship between reception and production. The term *productive noticing* is preferred. Here, it is noted that receptive ability is usually necessary to confirm the correctness of any production, but the feedback is indirect and the form to be *productively noticed* for learning remains the '*imitation*' itself: the manifestation of muscle positioning.

Removing *imitation* from the model explains the discrepancies of *imitation* not being derivable from receptive knowledge and of *imitation* not being necessary to develop receptive ability. Its precarious appendage was a symptom of attempting to view receptive and productive skills along a single process of development. Unfortunately, this provides only the model for developing receptive ability, with unnecessary appendages. *Productively noticing* the corresponding muscle position, as the result of having achieved the sound pattern by learning *how*, is the *beginning* of three<sup>5</sup> stages of developing productive ability. It can also be seen how receptive ability could aid the early stage of this *noticing*. Although the two inform each other via feedback loops, the direct productive development can only begin by *productively noticing*, or actually producing the muscle position making the form. This is summarised in Figure 2.

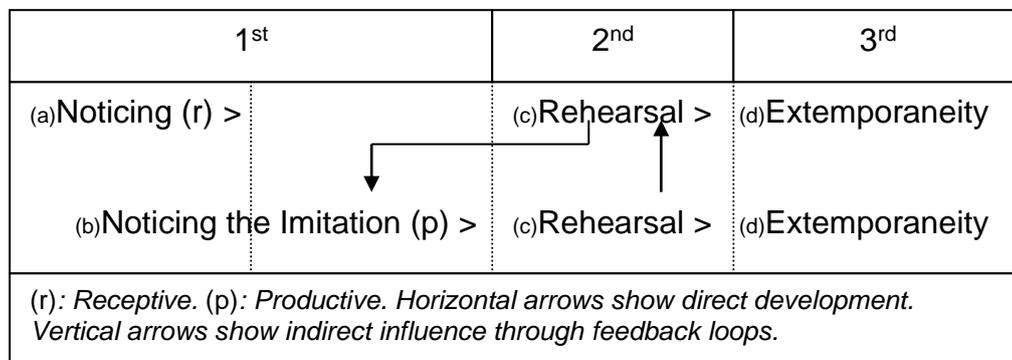


Fig. 2. Stages of development of the separate skills of receptive and productive ability.

The division of the *listening* category in the above survey slightly pre-empted this model. Without prejudice to any assertions made, the *de-contextual listening* serves *receptive noticing*, whilst the *contextual listening* serves *receptive rehearsal*. This reaffirms the assertion that prevalent methodology does develop receptive skill in an effective way. Affective factors notwithstanding, even Gimson's (1975) *identify*, followed by *listen* could, in theory, achieve these goals – producing them may additionally but diminutively *rehearse* any receptive ability through sensory feedback loops. This can be found in spin-off practice material such as Mortimer's (1985) *Element of Pronunciation* for supra-segmentals. Interestingly, Mortimer is not laden with technical explanations. They are not necessary for practicing receptive skills, nor do they correspond to acoustic manifestations. The technical explanations in Gimson and popular coursebooks are merely residual from the transfer of theory into practice, precisely mirroring and causing the unnecessary appendage of *imitation* into the prevalent model. This evidences the accidental nature of Gimson's (1975) and Morley's (1991) models, and provides a root for the misleading conflation of receptive and productive development.

This also reaffirms that the *repeat* activities attempt to *rehearse productively* forms that have not been produced at all, not yet *noticed* in the sense of muscle positioning. This initial stage is omitted in the prevalent model – except in the realm of segmentals, which, as acknowledged above, are taught with more satisfaction. Skipping the first stage, expecting learner to *productively rehearse* a form not yet *productively noticed*, is bound to failure.

### **Completing the Learning Process**

Although many models exist to complete the learning process, most notable Schmidt's application of Baars (Baars, 1988; Schmidt 1990), many build upon Kihlstrom's (1984) view of *rehearsal*. Being assumed for the purposes of brevity that *rehearsal*, in time, leads to *extemporaneity*: the ability to recall a form without concentration (Atkinson and Shiffrin, 1968; Kihlstrom, 1984; Baars, 1988; Schmidt, 1990). Importance is thus attached to *rehearsing* the correct form whenever concentrating, rather than reinforcing those incorrect.

*Productive rehearsal* proper only pertains through the sensory feedback loops when a form is produced or manifested through muscle positioning, matching with phonological knowledge through the auditory feedback loop. This slowly establishes *productive extemporaneity*. Without relying on chance, teaching must encapsulate *how* to pronounce unknown forms. The learner must learn to produce, *noticing* the muscle position of the form to *rehearse*.

### **Implementing *How Instruction***

The theory and experimental results do suggest that it is beneficial to centre exercises on *receptive noticing* and *receptive rehearsal*, listening to develop receptive skills. However, teaching demands do *not* stop here. Receptive ability is merely a prerequisite to learning *how* to produce a sound. *Productive rehearsal*, on the other hand, can only begin for those sound patterns learnt by chance and recognised as correct, unless *how* to produce is taught *before* such *rehearsal*. It appears that the *practicing* exemplified in the current literature, especially if integrated into a communicative classroom, is already well designed to fulfil the latter requirement as a second step in productive training.

As expounded above, the prevalent methodology does not adequately tackle the *how* problem in all areas of pronunciation. The first productive step of *noticing* appears to be neglected in many areas. Since it has been established that *receptive noticing* beforehand and *productive rehearsal* afterwards are both required, the current framework is left in place. Only *repeat* exercises are to suffer – possibly confined completely to testing purposes (Stevick, 1978) or translated entirely into communicative activities after *noticing*. Beyond this, instruction must be given in *how* to create sound patterns with which the learner is receptively familiar, to *notice productively*.

Although basic misconceptions in theory, which may initiate any developments in approach, are here considered, the implementation is in practice ever more complex. Implementation of approaches, or even syllabi or methodology for the classroom must not only take these amendments into account, but also centre any approach on the learners themselves within limited resources and in competition with other syllabi, such as grammatical and communicative, or other non-pronunciation components. They must also extend to supra-segmentals in addition to the already reasonably developed segmentals. The present focus and space not permitting, it falls elsewhere to develop any outline of a more complete approach. However, it does suffice to note in response to the present focus that without direct reference to the vocal apparatus, reliance is mainly placed on innate ability or chance, and once the learner has been instructed to *repeat* once or twice without avail, further commands would clearly become oppressive without further teaching, on *how* to create the sound. This teaching must refer directly tackle the question of *how*.

## Future Directions

In more critical consideration of future directions, the assumption has been maintained throughout that classroom teaching proceeds employing published material only. A number of references provide for the alternative possibility that teachers do a great deal more. Morley summarises ‘teacher-as-coach’ responsibilities, including points such as ‘[p]roviding models, cues, and suggestions for modifications of elements in the speech patterning for each student’ (1991, p. 508). This spans less than one page, tellingly without a single reference. The non-technical explanations in Kenworthy (1981), Kelly (2000), and elsewhere may be only a sample of teacher training in *making sounds* without formal planning. Even Derwing, Munro and Wiebe’s (1998) comment that teachers *have to* resort to intuition implies that teachers do *something* not found in publications. If classroom approaches differ significantly from those here analysed, a reappraisal of the evidence may be necessary. Thus, continuing engagement within pronunciation fora is endorsed for teachers to share techniques that bridge the gap between learners receptively familiarity with a form, and their producing it without reliance on chance.

## Conclusions

Despite apparent concordance in traditional approaches, signs of ineffectiveness and teacher dissatisfaction continue to accumulate. The analysis of the prevalent model of progression finds glaring discrepancies wholly indicative of this gap in pronunciation teaching between receptive familiarity and productive *rehearsal*. Many centred primarily on an *assumed* inter-relation between the development of receptive and productive skills, prompting the immature transfer of theory into the classroom.

The initial problem has been compounded by inserting a concept of *imitation* to maintain credibility with Kihlstrom's (1984) neurological model. This mistakenly rendered listening practice not *rehearsal*. Multiple additional dimensions of complication have arisen by entangling concepts of intelligibility with teaching approaches, shifting to onus onto the learner without instruction rather than describing the appropriate pronunciation target for the learner (an important but relatively unrelated question).

Combined with indications that teaching *how* to produce segmentals augments progression, it is suggested that this *instruction* be extended to all aspects of pronunciation. This includes, first and foremost, supra-segmentals. This is wholly consistent with, and even compliments, Kihlstrom's (1984) learning model in a more rational and consistent way. The more rational model, supported by considered research and empirical evidence from laboratories and classrooms, is better equipped to chart a more representative progression. This is mainly by separating the development of receptive and productive skills. It also reconciled the aforementioned discrepancies, placing receptive training and productive training alongside that of other linguistic forms. It also allows for supra-segmentals to be considered on par with segmentals, hopefully inducing more development in teaching materials for supra-segmentals. Some historical tracing locating patches attempting to amend previous misdirection outlines possible causes for the continuation of a now outdated model, which in turn proliferates incomplete teaching materials. At any rate, rethinking the model of progression will hopefully allow for more reflective applications of descriptive linguistics, mediated by theories of learning. Receptive stages must be seen to develop receptive ability, and productive stages to develop productive ability, possibly crossing over where conscious and justifiable.

For productive instruction, the gap between receptive familiarity with a form, and its *rehearsal*, leading ultimately to *extemporaneity*, must invoke new muscle positioning to create a new sound pattern. Reflecting the current uncommitted tendency for segmentals, an extension of utilising direct reference to the vocal organs and holistic approach to phonetics is necessary to plug the gaps in this area. For affective priorities, the teaching of *how* to create both segmental and supra-segmental forms must find less mundane outlets and foster learner collaboration in moving away from first language substitutes. It is beyond doubt that further research is necessary into the substance of techniques, and possibly into the nature of the particular difficulty perceived of or pertaining to supra-segmentals. Alternative theory here discussed may suggest that looking specifically to the currently lacking *how* instruction will go some way toward alleviating the burden of developing teaching techniques. Unfortunately, it is not within the ambit of the current paper to develop these requirements further. No doubt, however, more holistic approaches will demand reference to the vocal organs with dynamic, creative usage of the knowledge already in development – being essentially alike whether teaching segmental aspects, or supra-segmental aspects. Only then can *productive rehearsal* begin to learn the sound patterns for *extemporaneous* usage, granting pronunciation as accurate as the learner desires to aim.

## References.

Arteaga, D. L. (2000) Articulatory phonetics in the first-year spanish classroom, *Modern Language Journal*, 84(3), 339-354.

- Atkinson, R. C. & Shiffrin, R. M. (1968) Human memory: a proposed system, in K. W. Spence & J. T. Spence (Eds) *The psychology of learning and motivation*, vol 2 (London, Academic) 89-195.
- Baars, B. (1988) *A cognitive theory of consciousness* (Cambridge, CUP).
- Bongaerts, T., van Summeren, C., Planken, B. & Schils, E. (1997) Ultimate attainment in the pronunciation of a foreign language, *Studies in Second Language Acquisition*, 19(4), 447-465.
- Bongaerts, T. (1999) Ultimate attainment of L2 pronunciation: the case of very advanced late L2 learners, in D. Birdsong (Ed) *Individual differences and universals in language learning aptitude* (Rowley, MA, Newbury House) 133-160.
- Bowler, B. & Cunningham, S. (1999) *New headway: upper-intermediate pronunciation* (Oxford, OUP).
- Catford, J. C & Pisoni, D. B. (1970) Auditory vs articulatory training in exotic sounds, *Modern Language Journal*, 54(7), 477-481.
- Cauldwell, R. (2002) *Phonology for listening: relishing the messy*. Available online at: <http://bit.ly/2Aiujiwi> (accessed 21 November 2017).
- Chela-Fores, B. (2001) Pronunciation and language learning: an integrative approach, *IRAL: International Review of Applied Linguistics*, 39, 85-101.
- Cunningham, S. & Bowler, B. (2000). *New headway: intermediate pronunciation* (Oxford, OUP).
- Derwing, T. & Munro, M. J. (2005) Second language accent and pronunciation teaching: a research-based approach, *TESOL Quarterly*, 39(3), 379-397.
- Derwing, T. & Munro, M. J. (2009). Putting accent in its place: rethinking obstacles to communication. *Language Teaching*, 42(4), 476-490.
- Derwing, T., Munro, M. J. & Wiebe, G. (1997) Pronunciation instruction for fossilised learners, *Applied Language Learning*, 8, 217-235.
- Derwing, T., Munro, M. J. & Wiebe, G. (1998) Evidence in favour of a broad framework for pronunciation teaching, *Language Learning*, 48(3), 393-410.
- Elliot, A. R. (1995) Foreign language phonology: field independence, attitude, and the success of formal instruction in Spanish pronunciation, *Modern Language Journal*, 79(4), 530-542.
- Elliot, A. R. (1997) On the teaching and acquisition of pronunciation within a communicative approach, *Hispania*, 80(1), 95-108.
- Esling, J. H. & Wong, R. F. (1983) Voice quality settings and the teaching of pronunciation, *TESOL Quarterly*, 17(1), 89-95.
- Flege, J. E. (1981) The phonological basis of foreign accent: a hypothesis, *TESOL Quarterly*, 15(4), 443-455.
- Gardner, J. (1993) *Frames of mind*, 2<sup>nd</sup> Ed (New York, Basic).
- Gimson, A. C. (1975) *A practical course in phonetics* (London, Edward Arnold).
- Goto, H. (1971) Auditory perception by normal Japanese adults of the sounds 'L' and 'R', *Neuropsychologia*, 9, 317-323.
- Hahn, L. D. (2004) Primary stress and intelligibility: research to motivate the teaching of suprasegmentals, *TESOL Quarterly*, 38(2), 201-223.
- Hancock, M. (1995) *Pronunciation games* (Cambridge, CUP).
- Hancock, M. (2003) *English pronunciation in use* (Cambridge, CUP).
- Hewings, M. (2004) *Pronunciation practice activities* (Cambridge, CUP).
- Honikman, B. (1964) Articulatory settings, in D. Abercrombie, D.B. Fry, P.A.D. MacCarthy, N.C. Scott, & J.L.M. Trim (Eds) *In honour of Daniel Jones* (London, Longmans).
- Jenkins, J. (2000) *The phonology of English as an international language* (Oxford, OUP).
- Jones, D. (1956) *The pronunciation of English* (Cambridge, CUP).
- Kelly, G. (2000) *How to teach pronunciation* (Harlow, Pearson).
- Kenworthy, J. (1987) *Teaching English pronunciation* (London, Longman).
- Kihlstrom, J. F. (1984) Conscious, subconscious, unconscious: a cognitive perspective, in K. Bowers & D. Meichenbaum (Eds). *The unconscious reconsidered* (New York, Wiley) 149-211.

- Lane, H. & Schnieder, B. (1963) Methods for self-shaping echoic behaviour, *Modern Language Journal*, 47(4), 154-160.
- Laroy, C. (1995) *Pronunciation* (Oxford, OUP).
- Leather, J. (1983) Second-language pronunciation learning and teaching, *Language Teaching*, 16, 198-219.
- Levis, J. M. (1999) The pronunciation component in teaching English to speakers of other languages, *TESOL Quarterly*, 33(1), 37-63.
- Levis, J. M. (2005) Changing contexts and shifting paradigms in pronunciation teaching, *TESOL Quarterly*, 39(3), 269-377.
- Lord, G. (2005) (How) can we teach foreign language pronunciation, *Hispania*, 88(3), 557-567.
- Mann, V. A. (1986) Distinguishing universal and language-dependant levels of speech perception: evidence from Japanese listeners, *Cognition*, 24(3), 169-196.
- Marks, J. (2007) *English pronunciation in use: elementary* (Cambridge, CUP).
- Morley, J. (1991) The pronunciation component in teaching English to speakers of other languages, *TESOL Quarterly*, 25(3), 481-520.
- Mortimer, C. (1985) *Elements of pronunciation* (Cambridge, CUP).
- Mueller, T. H. & Niedzielski, H. (1967) The influence of discrimination training on pronunciation, *Modern Language Journal*, 52(7), 410-416.
- Neufield, G. G. (1980). On the adult's ability to acquire phonology, *TESOL Quarterly*, 14(3), 285-298.
- Neufield, G. G. (1988) Phonological asymmetry in second language learning and performance, *Language Learning*, 34(4), 531-559.
- Parish, C. (1977) Toward a practical philosophy of pronunciation, *TESOL Quarterly*, 11(3), 311-317.
- Pennington, M. C. & Richards, J.C. (1986) Pronunciation revisited, *TESOL Quarterly*, 20(2), 207-225.
- Pimsleur, P. (1963) Discrimination training in the teaching of french pronunciation, *Modern Language Journal*, 47(5), 199-203.
- Prator, C. (1971) Phonetics vs phonemics training in the ESL classroom, *TESOL Quarterly*, 5(1), 61-72.
- Oyama, S. (1976) A sensitive period for the acquisition of non-native phonological systems, *Journal of Psycholinguistic Research*, 5, 261-283.
- Riney, T. J., Takagi, N. & Inutsuka, K. (2005) Phonetic parameters and perceptual judgements of accent in English by American and Japanese listeners, *TESOL Quarterly*, 39(3), 441-466.
- Schmidt, R. W. (1990) The role of consciousness in second language learning, *Applied Linguistics*, 11(2), 129-158.
- Schnieder, W. & Chein, J. M. (2003) Controlled and automatic processing: behaviour, theory, and biological mechanisms, *Cognitive Science*, 27(3), 525-559.
- Setter, J. & Jenkins, J. (2005) Pronunciation, *Language Teaching*, 38, 1-17.
- Shiffrin, R. & Schneider, W. (1977) Controlled and automatic human information processing. ii. perceptual learning, automatic attending, and a general theory, *Psychological Review*, 84, 127-190.
- Stevick, E.W. (1978) Towards a practical philosophy of pronunciation: another view, *TESOL Quarterly*, 12(2), 145-150.
- Stevens, P. (1974) A rationale for teaching pronunciation, *ELT Journal*, 28(3), 182-189.
- Terrell, T. D. (1991) The role of grammar instruction in a communicative approach, *Modern Language Journal*, 75(1), 52-63.
- Timmis, I. (2002) Native speaker norms and international English, *Applied Linguistics*, 56(3), 240-249.
- Wennerstrom, A. (1994) Intonational meaning in English discourse: a study of non-native speakers, *Applied Linguistics*, 15(4), 399-420.
- Yule, G., Hoffman, P. & Damico, J. (1987) Paying attention to pronunciation: the role of self-monitoring in perception, *TESOL Quarterly*, 21(4), 765-768.

Yule, G. & MacDonald, D. (1995) The different effects of pronunciation teaching, *IRAL: International Review of Applied Linguistics*, 33, 345-350.

## Footnotes

- <sup>1</sup> It could alternatively be speculated that those with intelligences inclined towards acoustic, visual, or symbolic processing, for example, may store language in means also inclined towards that gift (see Gardner 1993). Perhaps it would be more accurate to suggest that central knowledge is not *necessarily* representative of a manifestation, and is not representative of all. Abstract knowledge has elsewhere been variously termed representations and adaptations.
- <sup>2</sup> Short term memory has variously been termed primary memory and working memory. Kihlstrom (1984, cited in Schmidt, 1990) suggests that, insofar as learning is concerned, short term store, consciousness and focal awareness are additionally interchangeable.
- <sup>3</sup> The *Practicing sound patterns* category only included tasks that would *not* constitute practice for any learner unable to produce the sound in question.
- <sup>4</sup> Three activities in Kenworth (1987) fell under two categories; as did two in Hancock (1996), four in Kelly (2000) and 11 in Hewings (2004). The frequency reflects this. When calculating percentages, these were counted twice. Again, minor inaccuracies are towards conservative since none of those classified as *making sounds* fell into a second category.
- <sup>5</sup> At first blush, proposing a three stage model may resemble Morley's (1991) model before noticing being appended to the beginning. It can be seen, however, that Morley clearly spoke about receptive training wherever envisaged. Thus one or more stages of the model would require division at some point to be reconciled with the division of receptive and productive skills. It is more likely Morley herself had receptive noticing, preceding production in mind.