Chapter 3
Defining prosody

3.1. Lexicographical prelude

Sometime around 2002 – the year in which the Speech Prosody conference series was launched in Aix-en-Provence – I happened to notice that the online Oxford English Dictionary (*OED*) still offered the same definition of *prosody* found in the first print edition of the 1930s. That original definition, omitting citations and a few irrelevant details, runs as follows:

> 1. The science of versification; that part of the study of language that deals with metrical composition; formerly reckoned as a part of grammar ..., and including also the study of the pronunciation of words (now called *phonology* or *phonetics*), esp. in relation to versification.
> [first citation ca. 1450]
>
> 2. Correct pronunciation of words; the utterance of the sounds of a language according to rule; observance of the laws of prosody. *rare*
> [first citation 1616].

The second edition of 1989, which is what I found online about 2002, keeps these first two points almost completely unchanged, and adds the following:

> 3. Linguistics. In the theories of J. R. Firth and his followers: a phonological feature having as its domain more than one segment.
> Prosodies include the class of 'suprasegmental' features such as intonation, stress, and juncture, but also some features which are regarded as 'segmental' in phonemic theory, e.g. palatalization, lip-rounding, nasalization.

Presumably nobody who set out for Aix-en-Provence in the spring of 2002 thought they were on their way to discuss versification or Firthian phonology. Yet nowhere did the *OED* give any indication that by 1989 (and certainly by 2002) a new meaning of *prosody* was not only already widespread, but had essentially superseded any earlier senses. It was only some time after 2002 – I don't know when – that the third definition in the online *OED* was changed to the following:

> 3. Phonology. A suprasegmental phonological feature such as intonation and stress. Also: such features collectively; the patterns of stress and intonation in a language.
> In early use sometimes (as in the work of J. R. Firth (1890–1960) and his followers) applied to some features which standard phonemic theory would regard as segmental, such as palatalization, lip-rounding, and nasalization [cross-reference to relevant senses of *prosodic* omitted].

Similar developments took place in many other European languages for the terms corresponding to *prosody* and *prosodic*: well into the 1990s, most dictionaries gave only traditional definitions related to the *OED*'s first two senses; after that, most record something like the meaning intended by the organizers of Speech Prosody.

3.2. The story of 'prosody'

The foregoing seems like a simple story of lexicographical inertia in the face of rapid developments of technical terminology. However, there is rather more to it than that. A brief sketch of the way the terms *prosody* and *prosodic* have been used suggests something almost like historical inevitability to the gradual shift in their meaning. At the very least, we see history repeating itself[1].

3.2.1. The classical background

If we take account of equivalent words in other European languages in addition to the English forms, we discover that the *OED*'s first sense of *prosody* – the science of versification – is not actually the term's original meaning. As *prosōdia* (προσῳδία), the term was used as early as the 6th century BC by Greek grammarians and philosophers, including Aristotle and Plato, to refer to the word accents of Classical Greek. This form is transparently composed of the prefix *pros-* (προσ-, with a meaning something like 'on' or 'to') and the root *ōdē* (ᾠδή) 'song', and therefore means something like the musical accompaniment – the 'song on top' – of the segmental sounds of a word. The centrality of the meaning 'word accent' is further shown by the fact that the Latin word *accentus*, like many Latin grammatical terms, is an early direct calque or loan-translation of Greek *prosōdia*: the combination of *ad* 'to' +*cantus* 'song' yields the form *accentus* by regular rules of phonological derivation. Note also that in this sense both Greek *prosōdia* and Latin *accentus* were regularly used in the plural as well as the singular; the word accents were 'prosodies', phonological events that occurred at a specific point in the word.

By the second century BC the term also referred to the written marks that were beginning to be used to indicate the accents of Classical Greek. (During the classical period itself – ca. 6th-4th centuries BC – accents were not marked.) Further extensions followed during the Byzantine period; by the second century AD, *prosōdia* was also used (in the writings of Sextus Empiricus) to refer to other phonemically distinctive properties that were not indicated in writing during the classical period but for which diacritical marks were later developed. In addition to word accents, these included vowel length (which had always been indicated by distinct letters for short and long /e/ and /o/ but was not indicated for /i a u/) and the presence or absence of initial aspiration (the difference between 'rough breathing' and 'smooth breathing'). In an even more extended sense, the term probably also referred to various notation schemes – forerunners of modern European punctuation – that were devised to help students of rhetoric speak more effectively from written text by indicating the natural groupings of words into phrases and larger units.

The shift of prosody's realm to the structure of verse – the *OED*'s first meaning – seems to have happened sometime during the Middle Ages, or perhaps when the word was imported into English and other European languages. It is at least possible that this shift was due to etymological confusion. There is another Ancient Greek word, *prosodios* (προσόδιος), meaning 'processional', which is based not on the root *ōdē* (ᾠδή) 'song' (with a long /o/), but on *hodos* (ὅδος) 'road' (with a short /o/). The

---

[1] Section 3.2 is based heavily on my reading of Allen 1973: 3-16 and Crystal 1969: 20-90, and on philological and bibliographical pointers from Nina Grønnum, John Joseph, Meg Laing and Henry Stevens. Any inaccuracies or misinterpretations are my own responsibility.

derived adjective *prosodiakos* (προσοδιακός) was applied to a type of metre suitable for processional songs, and to a type of foot characteristic of that metre. The Liddell and Scott *Greek-English Lexicon* records the forms *prosōdion* (προσώδιον) and *prosōdiakos* (προσωδιακός), but treats both as errors (*falsa lectio*) for the forms with the short /o/; the *OED* suggests that this confusion arose in Latin, which unlike Greek had only one letter for both short and long /o/. In other words, the fact that *prosodiakos* was used to describe a type of poetic metre may have led to a spurious link to *prosody*. Even if the shift to versification was ultimately based on confusion, though, it is beyond question that the meaning of *prosōdia* gradually expanded during the Classical and Byzantine periods. From an initial stage in which it referred only to Greek word accents, it was extended to refer to features of length and aspiration, and beyond that, probably, to features of grouping and phrasing.

3.2.2. Twentieth century linguistics

The gradual expansion of the classical terms is mirrored remarkably closely in the development of the words *prosody* and *prosodic*, and their counterparts in several other European languages, in the course of 20th century linguistics. Beginning early in the century, some linguists began to use these terms (particularly the adjectival form *prosodic*[2]) to refer to phoneme-like distinctions at the word level that are not conveyed by contrasts between phonetic segments – i.e. something like the original meaning of Greek *prosōdia*. The earliest instance of this usage of *prosodic* I have found in English is in a paper on Athabaskan relative clauses by Sapir (1923: 137), where he states that two forms may be distinguished by 'a prosodic difference (one of stress or pitch)'. About the same time, one of Sapir's followers (de Angulo 1929: 117) proposed *prosody* as a cover term for the application of such distinctions in morphological processes; this proposal covers quantity (e.g. vowel length distinctions) as well . Similar uses, especially of *prosodic*, are found in the writing of others in Sapir's circle, notably Morris Swadesh (e.g. 1934, 1949), and are subsequently found in early work by Trager and Bloch (1941). However, Trager and Bloch tended to use *suprasegmental* in later writing with approximately the same meaning, and this became standard Bloomfieldian terminology. As for other languages, the 1933 edition of Marouzeau's *Lexique de la Terminologie Linguistique*, which records technical usage in French, German and English, contains an entry for the noun *prosodie* which gives only the classical Greek meaning and the meaning related to metrics and versification; by the time of the 1943 edition, the entry has been revised to add that 'phonologists' use the term to refer to 'dynamic, melodic, quantity-related, etc.' [my translation] phonetic properties of a language. The German adjective *prosodisch* is used without comment by Trubetzkoy throughout *Principles*, especially in chapter IV section 5, referring primarily to what he also calls 'rhythmic-melodic' features; for Trubetzkoy, prosodic features definitely include distinctive quantity.

---

[2] There is a definite asymmetry between the noun and the adjective in the literature. Though the adjectival use ('prosodic features', etc.) can found by the early 1940s in the work of a variety of prominent linguistic scholars, the noun seems to have retained the primary sense of 'the science of versification' in many European languages for much longer, and other linguistic uses remained rare until the 1970s. Perhaps in a linguistic context the noun *prosody* seemed at greater risk of creating ambiguity, whereas the adjective *prosodic* could be combined with other nouns in such a way as to avoid suggesting the poetic meaning.

In short, the terms *prosodic* and (to a lesser extent) *prosody*, and their equivalents, were already widely known in both European and American linguistics by the early 1940s, generally referring to word-level distinctions of tone and accent and, for at least some writers, quantity. Before long they were applied to phonetic phenomena beyond the level of the word as well, in particular to phenomena involved in grouping words into phrases, phrases into utterances, and so on. As early as 1934 Swadesh spoke of 'sentence prosody' (1934: 122), while Trager and Bloch (1941) were among the first to mention 'juncture' (phonetic cues to boundaries) in connection with stress and pitch. Further expansion of the linguistic meaning followed in the 1950s, notably in the context of clinical work on aphasia and other speech and language disorders, and psychological and psychiatric work dealing with the expression of speaker affect and personality. Researchers in these fields (and more recently, researchers concerned with technological applications such as automatic speech recognition and synthesis) have long taken for granted that there is a fundamental distinction between propositional content and everything else conveyed by speech. This distinction, which Crystal (1969: 76) characterizes as 'verbal' vs. 'vocal', goes back to the early 20th century (see Crystal 1969: 62-90 for an extensive review) and is still a major driver of research today (for reviews see e.g. Frick 1985, Scherer et al. 2003, Belin et al. 2011 on the psychological side; Schröder 2001, Shriberg and Stolcke 2004 on the technological side). The use of the terms *prosody* and *prosodic* for the non-verbal side of this distinction seems to date roughly from the middle of the 20th century. An early instance of this usage – perhaps its source – is a still-cited article on 'dysprosody' by Georg Monrad-Krohn (1947). Monrad-Krohn's paper was the first scientifically respectable report of 'foreign accent syndrome', and identified disturbances of speech rhythm and melody, rather than segmental misarticulation, as a major contributor to the perceived foreign accent.

Nevertheless, presumably because of competition or potential confusion with the traditional meaning referring to metrics and versification, and (in North America at least) because of the availability of the effectively synonymous term *suprasegmental*, linguistic use of *prosody* and *prosodic* remained relatively limited until the 1970s. In English, it is possible that the expansion of the linguistic meaning was further inhibited by the idiosyncratic application of the terms *prosody* and *prosodic* in the work of J. R. Firth, whose 'London school' was extremely influential in British linguistics from the 1930s to the 1970s. Firth (e.g.1948) rejected the strongly segmental basis of the phoneme idealization as it had developed in the 1930s and 1940s, and took seriously the idea that many phonetic properties apply to stretches of speech longer than a single segment. He designated any such property as a 'prosody', using the term as a noun with a plural as in Classical Greek (to the consternation, many years later, of Microsoft's grammar-checker). Firthian phonological representations were a complex mix of prosodies and what were called 'phonematic units' (segment-size clusters of irreducibly local phonetic properties), and the Firthian approach to phonology was known for a time as 'prosodic analysis' or 'prosodic phonology'. (For summaries of Firthian phonology see Anderson 1985 or Ogden and Local 1994; for examples see several of the papers in Palmer 1970, especially Henderson 1949.) Firth and his followers were especially interested in phenomena like vowel harmony and nasal spreading, which provide an obvious justification for positing abstract phonological units that are linked to domains larger than segments. Some of Firth's ideas were revived or rediscovered and developed in the descriptive work of the 1970s and 1980s in the tradition of autosegmental phonology (see chapter

1), but they played no direct role in the rapid shift in the meaning of *prosody* during the 1970s, and the specifically Firthian usage is appropriately treated by the *OED*'s latest definition as no longer current.

It was not until Halle and Keyser first published their work on English metre (1966), triggering lively scholarly debates (e.g. Beaver 1968, Keyser 1969, Sledd 1969, Wimsatt 1970, Standop 1972), that the linguistic usage of 'prosody' finally began to come into its own. By relating facts about poetry to linguistic ideas about the phonological structure of utterances, Halle and Keyser's work suggested common themes between prosody's new linguistic meaning and its traditional poetic sense. The mingling of phonological and poetic concerns was taken further in Mark Liberman's doctoral thesis (1975), which drew explicit links between musical text-setting and what he called 'tune-text association' in intonation. Liberman's pioneering work drew the field's attention to the importance of hierarchical structure for the analysis of intonation and stress, and by the 1980s an enormous amount of work was being carried out under the rubrics of 'metrical phonology' and 'prosodic phonology'[3]. Several major works in this tradition appeared during this period, including Selkirk 1984, Nespor and Vogel 1986, and Pierrehumbert and Beckman 1988, all of them focusing in various ways on the role played by hierarchical structure in phonology. Importantly, the phenomena treated in these works included not only intonation and stress, but also phonological boundary effects such as external sandhi and Trager-Bloch 'juncture'. Quantitative bibliographical evidence, based on a rough analysis of entries in Google Scholar™ carried out in May 2011 (details in section 3.6 below), clearly shows that a significant shift in usage began about the time of Halle and Keyser's work, and that the poetic sense of *prosody* and *prosodic* was virtually obsolete by about 1985.

One possible objection to the foregoing summary is that the broader linguistic sense of the terms may already have been current in the Scandinavian languages before the 20th century. Both *prosodi* and *prosodisk* are used extensively in the works of the Swedish grammarian Adolf Noreen, notably in his major work *Vårt Språk* ('Our Language'), which was published in several volumes beginning in 1903. Under the heading of *prosodi* Noreen discusses features of intensity or prominence, melody and quantity, as well as syllable structure and phonotactics. According to Jakobson and Waugh (1979:142f), Noreen is the originator of the idea of a fundamental distinction between 'prosodic' and 'inherent' phonological features, which Jakobson developed in his own work (see further section 3.4.4 below). Whether Noreen's innovation lay partly in the use of the actual term *prosodi* or strictly in the drawing of the distinction is not clear from Jakobson and Waugh's discussion. However, it is at least possible that something like his use of the term was already familiar to Scandinavian scholars, who in their own languages were acquainted with word-level phenomena not unlike the Classical Greek accents. The Danish grammarian Jens Høysgaard, cited in the great 20th-century Danish dictionary *Ordbog over det Danske Sprog*, used the term *prosodi* in 1769 to refer to stress and quantity in words, but it is unclear whether he intended this as a departure from the metrical/poetic usage. It may or may not be coincidence that Monrad-Krohn, the one who coined the term *dysprosody* to describe the features of speech affected in foreign accent syndrome, was a speaker of Norwegian; this may have seemed a natural extension of an established Scandinavian

---

[3] Not to be confused with the use of 'prosodic phonology' to refer to Firthian descriptive work!

usage. However, even if this usage was already familiar in the Scandinavian languages, it seems fairly clear that it was not current in English, French, or German before about 1920.

3.3. 'Prosody' as miscellany

In less than a century, then, the modern linguistic terms *prosody* and *prosodic* appear to have undergone a development that is parallel to that of Ancient Greek *prosōdia.* Starting from a specific reference to word-level tonal or accentual features, both the classical and the modern terms expanded to include a range of other phenomena not normally indicated in writing, including in particular features of quantity and cues to the grouping of words into phrases. One possible interpretation of this parallel is that modern scholars have rediscovered a valuable insight into the organization of phonology, and that there is some natural unity to the range of things grouped together under the expanded meaning of the terms. Another is that the expansion of the meaning simply reflects the biases induced in classical times by alphabetic literacy and in the modern era by IPA transcription: anything not written with consonant and vowel letters must by definition be something else.

3.3.1. An alphabetic artefact?

On the face of it, there is a good case for regarding the parallel expansions of *prosōdia* and of *prosody* as an artefact of alphabetic literacy. The segmental idealization that underlies any type of alphabetic writing, including IPA transcription (see chapter 2), is intrinsically poorly adapted to representing certain phonological features that are generally taken to be 'prosodic'. This is because alphabetic representations are ordered strings of atomic elements, and as such they are ill suited to representing various readily perceptible phonetic properties of speech.

First consider duration. Alphabetic writing effectively excludes any indication of actual time: the only temporal property that counts, formally speaking, is linear order or precedence. There are ways of indicating categorical phonemic distinctions like that between long and short vowels, including diacritic marks (e.g. IPA [ː] or the acute accent in Hungarian orthography), diacritical letters (e.g. the letter <h> following a vowel letter in German orthography), and the device of writing two adjacent identical symbols (e.g. Finnish orthography). But these devices cannot readily be used to represent gradiently variable temporal features that have other linguistic functions, such as cues to stress and phrasing, nor can they reflect segment-specific durational properties like the fact that fricatives are generally longer than stops. These features are either subsumed under transcriptional abstractions like boundary symbols in IPA transcription and punctuation in ordinary alphabetic writing, or are simply ignored. Note in this connection that the IPA boundary symbols, including the symbol for 'linking (absence of a break)', are listed under the heading 'Suprasegmentals'.

Now consider pitch. A different consequence of the fact that alphabetic writing involves ordered strings is that it has difficulty representing phonological distinctions that are based on the overlapping or otherwise unordered arrangement of distinct phonological elements. If we are restricted to a string of symbols in which the linear order of the symbols represents succession in time, where do we put the symbol

representing a phoneme (such as tone) that occurs concurrently with another phoneme or phonemes?[4] In some alphabetic writing systems applied to tone languages (e.g. Dinka), the answer is simply that we don't put it anywhere; in others (notably Vietnamese), we use diacritic marks. Either way, tone thereby appears to be a phenomenon apart.

The impression that prosody is little more than the residue of segmental transcription is strengthened by the near synonymy of the terms *prosodic* and *suprasegmental*. As we saw earlier, Bloomfieldian linguists like Trager and Bloch adopted the latter term – which transparently takes the segmental idealization of phonetics as a starting point – in preference to the former. Throughout the 1950s and into the 1960s *suprasegmental* was the normal way of referring collectively to features of pitch, stress, and duration, at least in North America, and the term is still reasonably common – especially in definitions of *prosody*! The virtual interchangeability of the terms can be seen in Lehiste's book *Suprasegmentals*, which opens with the phrase 'The study of prosody' (1970: 1). At the very least, it is not implausible to claim that the segmental idealization of phonetics gives rise to the widespread view that the suprasegmental features represent a coherent set of phenomena.

Of course, there may be other reasons why suprasegmental distinctions are often excluded from alphabetic writing. One is the historical circumstance that alphabetic writing grew up in a part of the world where phonemic distinctions of tone are unusual. Another possible alternative reason is the articulatory basis of suprasegmental distinctions. The supralaryngeal gestures that give rise to vowels and especially consonants are fairly accessible to proprioception and observation of one's own movements, as anyone who has taught practical phonetics knows. It is easy to become aware of the articulatory difference between, say, [m] and [n], and hence of the need to provide distinct alphabetic symbols for distinct phonemes based on that difference. It is less easy to become aware of the way in which we produce phonemic distinctions of pitch and accent, and therefore perhaps less obvious that such distinctions should be represented.

Nevertheless, it is clearly true that the pitch and duration-related phonetic features that are at the core of most implicit definitions of prosody are difficult to represent given certain formal properties of alphabetic writing. The fact that the same features may be intrinsically more difficult to observe in one's own speech merely further hampers their incorporation into writing of any sort. We cannot exclude the possibility that the phenomena conventionally grouped together under expanded definitions of prosody

---

[4] There is no non-arbitrary answer to this question, as can be seen from the practice of alphabetically-oriented linguists describing the phonology of tone languages. In the collection of chapter-length descriptions of Sino-Tibetan languages in Thurgood and LaPolla (2003), most of the authors of the individual chapters have occasion to describe the 'syllable template' or 'syllable canon' of the languages they are describing, which in most cases involve lexical tone. A few authors (e.g. Bradley on Lisu, chapter 14) give formulas like C(G)VT [i.e. consonant, optional glide, vowel, tone] with the tone linearized at the end, after the specification of the segmental positions. One author (Mazaudon on Tamang, chapter 18) gives a similar formula, but with tone linearized at the beginning, before the segments, while yet another author (Solnit on Eastern Kayah Li, chapter 38) writes C1(C2)(G)V/T, explicitly using the / between the V and T symbols to 'reflect simultaneous occurrence'. In two chapters (Bauer and Matthews on Cantonese, chapter 9, and Wiersma on Yunnan Bai, chapter 40) we find multi-linear representations. The most common solution is simply to give formulas showing only the segments (e.g. (C)(G)V(G) in Gong's chapter on Tangut, chapter 37) and discuss tone separately.

are grouped that way primarily because alphabetic writing and IPA transcription have shaped our intuitions about them, not because they actually share essential properties in spoken language itself.

3.3.2. Definitions and lists

The suspicion that the extended linguistic sense of prosody does not actually refer to a coherent group of phenomena is strengthened when we consider attempts to provide a definition of the term. The problem was stated clearly by Lehiste (1970: 1f):

> Yet a certain degree of vagueness seems to characterize most discussions of prosodic features. They seem more elusive than segmental features, and their incorporation into a linguistic system sometimes seems to strain the limits of an otherwise coherent framework.
>
> This vagueness extends to the definition of prosodic features. In American linguistics, the term is used more or less synonymously with suprasegmental features. Suprasegmental features are usually either listed as the set of features consisting of pitch, stress, and quantity, or defined as features whose domain extends over more than one segment (Hamp 1957). A definition is preferable to a list; the definitions referred to, however, have at least two weaknesses. … If it is true that stress, pitch, and quantity behave in a way that sets them apart from features determining segmental phonetic quality, the definition should be revised.

This is not a new problem: Adolf Noreen, in introducing his ideas about the nature of prosody mentioned above, had this to say (Noreen 1903-07: 406, my translation[5]):

> Hitherto the term prosody (prosodic) has been taken in such a limited sense that it has merely designated the quantity, intensity and tonality of sounds, or indeed sometimes only their quantity, but for such a restriction there is no well-founded reason.

And Lehiste's remarks are equally applicable today, as can be seen from more recent attempts to define *prosody* in technical sources. For example, Crystal's brief definition in the glossary that forms part of the Oxford *International Encyclopedia of Linguistics* (Crystal 1992: 328) runs as follows:

> **prosody**: Variation in pitch, loudness, tempo, and rhythm, as encountered in any use of spoken language (thus subsuming the traditional sense of the metrical features of versification); also called **prosodic features**, and in phonemics analyses in terms of **prosodemes**. In generative phonology, prosodic features are one of the main dimensions of speech sound classification. In Metrical Phonology, one of the levels of structure in a metrical tree is the **prosodic level** [*sic*]. The canonical pattern of segments in a form is a **prosodic template**. In **Prosodic Phonology**, a **prosody** is a

---

[5] Disclaimer: I decipher Swedish rather than reading it. Thanks to Merle Horne for helping me track down and make sense of Noreen's work

feature extending over a stretch of utterance (such as the syllable of sentence), contrasting with the segmental notion of phonematic units.

Equally revealing of the difficulty of identifying the core of prosody is the dilemma experienced by ordinary dictionary makers. As I stated in the 'lexicographical prelude' above (section 3.1), beginning in the 1990s dictionary makers across the Western European languages suddenly took note of the changed meaning of the terms corresponding to *prosody* and *prosodic*. But a look at their new definitions reveals that they too largely fall back on providing illustrative lists of specific topics. The *OED*, as we saw earlier, now defines prosodic features as 'suprasegmental' and exemplifies that term with intonation and stress. De Mauro's linguistic definition of Italian *prosodia* also makes reference to suprasegmental phenomena, specifying them as 'intonation, intensity and duration'. The second edition of the Robert French dictionary, like the first edition, relies heavily on the passage from Marouzeau's technical *Lexique* quoted earlier, referring to the 'dynamic, melodic, quantity-related, etc.' phonetic properties of a language.

Furthermore, if we compare the lists provided by the different dictionary makers, we see notable differences of emphasis. The Seco et al. Spanish dictionary gives two separate sub-definitions of the modern sense of *prosodía*, one an attempted definition (in terms of phonetic features whose domain is larger than the phoneme, as in the passage from Lehiste just quoted), and the other a list specifying features of intonation and accent. The definition of *prosódia* in the Academy Portuguese dictionary is similar, but gives an even longer list of things that exemplify it: 'tone, intonation, rate, pause, accent, rhythm, intensity, etc.' The linguistics-related definitions of German *Prosodie* in both Wahrig and Duden, unlike those in the other languages, emphasize prosody's chunking function, i.e. its role in the division of the stream of speech into words and phrases. Again, though, they give accent and intonation as examples.

3.4. Distinctions that may be relevant to the definition of prosody

The sheer diversity of dictionary makers' attempts to get to grips with the linguistic meaning of *prosody* provides a glimpse of the theoretical confusion that lies behind its 20th century expansion. Yet there are a number of ways in which some coherent notion of prosody might be motivated, and beginning with Trubetzkoy we find explicit attempts to provide a sound theoretical basis for distinguishing prosodic features from other phonetic and phonological phenomena. These are reviewed in this section, in roughly chronological order.

3.4.1. Source vs. filter

One possible definition of *prosody* and *prosodic* is based on the distinction between 'source' and 'filter', in the now widely-used sense of those terms based on Fant's acoustic theory of speech production (1960). The speech signal, in this conception, is the result of passing a source of acoustic energy (most often, the pulse train emanating from the larynx) through a filter (the variously shaped supralaryngeal tract, which modifies the spectrum of the source signal). The idea of defining prosody on this basis was actually suggested by Trubetzkoy in *Principles*:

> The process of phonation of human speech can best be illustrated by the following scheme: somebody whistles or sings a melody into the mouthpiece of a tube and alternately opens and covers the other end of that tube with his hand. It is clear that three types of elements can be distinguished acoustically in the course of this process: first, the segments between closing and opening the orifice; second, the segments between opening and closing it; and third, the segments of the melody whistled or sung into the tube. Elements of the first type correspond to consonants, elements of the second types to vowels, and those of the third type to prosodic units. (1969: 93f).

However, Trubetzkoy contradicts this definition only a few pages later when he offers 'rhythmic-melodic' as a synonym for *prosodic*, and in much of what he says about prosody it seems clear that he is primarily concerned with pitch and duration.

Why is this a contradiction? A moment's thought makes clear that 'rhythmic' and 'melodic' properties are quite separate: the melody, in the sense of the pitch contour of the output signal, is indeed a function of the source, but durational and (more broadly) rhythmic patterns are not. In terms of Trubetzkoy's tube analogy, rhythmic properties of the signal do not depend primarily on what is sung or whistled into one end of the tube (the source features), but rather on the temporal details of how the other end of the tube is opened and closed by the hand (the consonantal and vocalic modifications of the supralaryngeal filter). That is, the perceived rhythm of speech is determined in great measure by such parameters as the proportion of vocalic to consonantal segments, the variability of the duration of the vocalic segments, and so on. This general understanding of speech rhythm was first clearly articulated by Dauer (1983) and is the basis of several recent attempts to quantify speech rhythm on the basis of such parameters (Ramus et al. 1999, Low et al. 2000, White and Mattys 2007, and cf. Arvaniti 2012). It is also related to MacNeilage's ideas (e.g. 1998) about the relationship between basic syllable rhythm and the natural physical periodicity of the opening and closing of the jaw. If we want to consider rhythm to be part of prosody, then we cannot distinguish prosodic features from other phonological phenomena on the basis of a distinction between source and filter. Conversely, if we want to pursue a definition of prosody in terms of source features, we will probably need to exclude rhythm.

3.4.2 Non-verbal vs. Verbal

Another conceivable basis for distinguishing prosodic from other features is to relate it to the difference between propositional content expressed in words and everything else conveyed by speech. As noted in section 3.2.2 above, this very broad understanding of what *prosody* involves is especially common in psychology and psychiatry, in speech therapy, and in speech technology. There is almost certainly a valid distinction to be drawn along these lines (this is the topic of chapter 4), though I greatly prefer the term *paralinguistic* for most of what is subsumed under prosody when it is defined in this way . There do seem to be universal aspects of the way some characteristics of individual speakers are conveyed by speech, the most conspicuous being the biologically-based differences between adult male and female voices. Manifestations of emotional state covered by informal terms like 'raising one's voice' probably also have a biological basis and can also be included here, though there are well-known differences among cultures regarding the social

acceptability of displaying these manifestations. The production and perception of these features can be impaired by brain damage or psychological disturbances, largely or perhaps even entirely independently of language, and it seems clear that the clinical/psychological focus on the non-verbal aspects of spoken communication is based on a genuine distinction.

Yet by identifying these aspects with 'prosody', we immediately have to grapple with the contradiction between expecting prosody to be related to the expression of emotion and attitude, on the one hand, and assuming that anything non-segmental is prosodic, on the other. The most obvious problem is that some non-segmental features, such as lexical tone, manifestly do contribute to propositional content, and are presumably unrelated to the expression of emotion. Even in non-tonal languages there are cases like the following pair (Rooth 1985), in which the placement of sentence accent creates two distinct meanings with different truth conditions:

(3.1) i. John only introduced Bill to SUE.
      ii. John only introduced BILL to Sue.

By and large there are enough empirical research questions on non-verbal communication that psychiatrists and social psychologists and even speech technologists may be forgiven for not thinking about this theoretical contradiction. For linguistics, resolving the contradiction has a much higher priority. One line of attempts to do this, for example, has involved researchers in looking for emotion-related explanations for intonation (e.g. Bolinger 1972, 1986). In my view, these are generally pretty implausible and (worse) untestable (Ladd 1987), but they may nevertheless be valid. Simply ignoring the whole problem, though, effectively means viewing lexical tone as something fundamentally exotic and mysterious (e.g. Martinet 1980: 83ff.) – an inappropriately ethnocentric view for a field that aims to understand both the unity and the diversity of language.

3.4.3 Suprasegmental vs. segmental

The success of segmental representations of speech in alphabetic writing generally, and in IPA transcription in particular, makes it tempting to see anything that cannot be written in terms of segments as part of a distinct class of phenomena. As we already saw, this implicit basis for defining 'prosody' seems to have developed independently in the classical Graeco-Roman world and in 20th century linguistics. The theoretical basis of the distinction, of course, has never been linked explicitly to alphabetic representation, but the very term 'suprasegmental' – adopted in the early 1940s by e.g. Trager and Bloch (1941: 224) and Hockett (1942: 8) – suggests that these features are somehow 'above' the string of segments and therefore presupposes a segmental phonetic idealization of the stream of speech (see chapter 2, section 2.3.4).

The two notions that figure most prominently in actually defining 'suprasegmental' are that suprasegmental features (a) occur simultaneously with segments and (b) extend over a longer domain than a single segment. Simultaneity seems to be an important part of what led to the identification of pitch, stress and quantity (e.g. Trubetzkoy's 'rhythmic-melodic features') as the essential components of prosody. The segment is defined by a limited set of phonetic properties – manner, place, voicing, and so on – and any phonetic properties that are left over must be seen as

something separate, occurring simultaneously with the segment but not part of what defines it. If there were a principled basis for determining which phonetic properties are segmental, then this definition of suprasegmental would be defensible, but insofar as the definitions of segmental properties are based on Eurocentric alphabetic assumptions, the definition becomes rather circular. There are many ways, for example, in which pitch and voicing (or perhaps more correctly, fundamental frequency and voice onset time) interact in the identification of 'segmental' distinctions, as we saw in the discussion of Kera in chapter 2, section 2.3.2. A currently well-studied case of this sort involves the three-way laryngeal contrast in Korean stop consonants (e.g. Silva 2006, J. Kirby 2013).

As for the idea that suprasegmental features have intrinsically greater temporal extent (or, more abstractly, that they apply to larger domains in the phonology), it appears slightly less circular, in that it depends only on the existence of some segmented phonetic representation, not a specific set of 'segmental' properties. Here, too, however, there are conspicuous problems. One such problem is that at least some of the features that are commonly regarded as suprasegmental do not necessarily extend over multiple segments. This is true almost by definition for quantity distinctions, which mostly involve phonological properties of individual segments. Lexical tone can also pose problems here: in a syllable consisting of a voiceless consonant, a vowel, and a tone, the tone applies to exactly one segment. This makes it more difficult to justify separating the tone out from the other phonetic properties of the vowel segment. This problem was acknowledged in the 1940s by Zellig Harris, who commented on the arbitrariness of treating tone as a component separate from a vowel and suggested that 'we could just as well state that a language has not, say, 5 vowels and 3 tones, but 15 vowel phonemes' (1944: 200 [1966: 135; 1972: 129]).

The opposite problem is also widespread, and is in some sense more fundamental: phonetic features other than rhythmic-melodic ones often do spread themselves over multiple segments in sequence. If we decide to treat pitch differently on the basis that it applies to more than one phonetic segment in sequence, logic suggests that the same treatment should apply to features of place or voicing or nasality when they apply to two or three successive segments. (For example, since English obstruent clusters in syllable codas invariably agree in voicing (e.g. *cats* [kæts] has a voiceless final cluster vs. *cads* [kædz] has a voiced one, but *[kætz] and *[kæds] are impossible), it seems justifiable to treat voicing as a suprasegmental feature spanning the entire syllable coda.) Considerations of this sort lay at the root of some of Hockett's early discussions of phonetic features (e.g. Hockett 1942) and Zellig Harris's much more thoroughgoing exploration of the notion of 'long components' in phonology (Harris 1944). However, the only school of thought to take such matters seriously was Firthian prosodic phonology. As we saw in section 3.2.2 above, Firthian theory did not ascribe any special status to rhythmic-melodic features, and insisted on defining 'prosodies' in language-specific terms. Features that regularly characterize domains longer than a single segment were treated as prosodies of those specific domains (e.g. 'syllable prosodies'), and considerable theoretical attention was paid to defining the kinds of phonological structures that constitute the domains to which prosodies can apply. Most other phonological theorists, though, have acknowledged the logical puzzle but have never allowed it to override the intuitive sense that there is a difference between (say) pitch and place of articulation. Except in Firthian linguistics, the assumption that pitch, stress and quantity are a coherent set took

precedence over other considerations, and proposals like Harris's long components made little theoretical headway.

3.4.4. Prosodic vs. inherent

Perhaps the most serious and credible attempt to provide a motivation for treating the rhythmic-melodic features as a coherent class rather than as an assortment of phonological leftovers was made by Roman Jakobson. The basic idea is clearly stated by Jakobson, Fant and Halle (1952; here referred to as *JFH*) in the context of their fully elaborated theory of distinctive features, which involves a fundamental distinction between 'inherent' and 'prosodic' features. The inherent features are the familiar features that distinguish one segment from another – consonantal, vocalic, strident, grave, and so on – while the prosodic features are apparently limited to those of pitch, stress, quantity and perhaps syllabicity (the presentation in *JFH* is brief and rather unclear on this point). The defining characteristic of the two types of features, according to *JFH* (p. 13), is that the inherent features can be identified at a single time-slice of the signal 'without any reference to the sequence', while the prosodic features 'can be defined only with reference to a time series'. On this view, a high vowel or a labial stop can be recognized as such without reference to any other part of the signal, but a stressed syllable is stressed only by comparison with an unstressed syllable, and a high tone is high only by comparison with a low-toned syllable. Jakobson seems to have held to this idea throughout his life; it is elaborated further in Jakobson and Waugh 1979, where (as noted in section 3.2.2 above) the original statement of the idea is attributed to Noreen (1903-07)[6]. The same approach to defining suprasegmental phenomena as a coherent class was adopted by Lehiste in her book *Suprasegmentals* (1970: 35f).

However, this definition is less successful than it first appears. It is not difficult to think of 'inherent' features that require reference to a time series: cues to the place of articulation of both oral and nasal stops are well known to reside primarily in the adjacent vowels, and the distinction between an unaspirated and an aspirated voiceless stop is partly a matter of what happens between the closure and the onset of the following vowel. Even the correct perception of vowel quality depends to some extent on having a conception of the speaker's vowel space, which is derived in part from other vowels in the same utterance (Ladefoged and Broadbent 1957; Nearey 1989). Conversely, the idea that 'prosodic' features invariably require within-utterance comparison is difficult to maintain, especially for pitch: it is perfectly possible in many African languages to have sentences consisting of e.g. only high-toned syllables, which can be identified as such even without the presence of other tones for explicit comparison. I have discussed the problem of pitch normalization at greater length elsewhere (Ladd 2008a: 188-210 [1996: 252-269]), arguing for a phonetic notion of 'tonal space' entirely analogous to that of 'vowel space'. The only important difference between pitch and vowel quality in this respect is that the former may vary more conspicuously from one speaker to another or from one occasion to

---

[6] What Noreen actually said (1903-07: 406) is that prosodic features are 'those that emerge when comparing sounds to each other in connected speech' [*de som hos ljuden i sammanhängande tal framträda vid deras inbördes jämförelse*] . The word I translate here as 'features' is *egenskaper*, which is directly cognate with German *Eigenschaften* – i.e. features in the sense of 'properties' or 'attributes' (cf. chapter 1, section 1.2.1, and chapter 2, footnote 5). Again, thanks to Merle Horne for help in dealing with Noreen's original text.

another. In neither case is there any acoustic invariance across speakers: some calibration to the speaker is always necessary. There may be justification for an idealization in which some features are instantaneous while others are intrinsically based on a comparison between two points in time, but it is difficult to base such an idealization on experimental evidence from speech perception, and difficult to make it neatly distinguish pitch, stress and quantity from everything else.

3.4.5. Syntagmatic vs. paradigmatic

Despite the objections just summarized, Jakobson's idea (or perhaps it is Noreen's) comes close to capturing the essence of an important phonological distinction. There really is a fundamental divide between syntagmatic features – ones that intrinsically involve a comparison between one element and another within the phonological string – and paradigmatic features – ones that involve a choice from a language-specific set of alternatives at a given point in structure. However, the Jakobsonian version of this idea is undermined by two interrelated misconceptions. First, it takes the basic insight and tries to make it compatible with a preconceived idea of which phonological features are prosodic, namely Trubetzkoy's 'rhythmic-melodic' features. Second, because its notion of syntagmatic comparison is anchored in real time (and, arguably, in speech perception) rather than abstract linguistic structures, it fails to acknowledge that there are actually significant differences between broadly 'melodic' features (which are functionally and structurally quite similar to segmental features) and broadly 'rhythmic' features (which really do involve syntagmatic comparison as part of their phonological essence).

The idea that stress distinctions are intrinsically syntagmatic was a major contribution of Mark Liberman's doctoral thesis (Liberman 1975 [1979]; Liberman and Prince 1977). According to Liberman, any pair of sister nodes in a hierarchical phonological constituent structure must be in a prosodic relation of prominence in which one is subordinate to the other. For example, in a simple phrase like *five pounds*, there are two possible relations, weak-strong and strong-weak, which may be diagrammed as follows:

(3.2)
```
           w    s                    s    w
   a.      five pounds          b.       five pounds
```
Such prominence relations apply at all levels of the structural hierarchy, as in the phrase *seven dollars*:

(3.3)
```
          w    s                       s    w
        s w  s w                     s w  s w
   a.    seven dollars       b.      seven dollars
```

(The (a) versions are pragmatically appropriate in a wide variety of circumstances, whereas those in (b), informally speaking, have 'contrastive stress' on the number.) The complete picture put forth in Liberman's analysis does involve local paradigmatic elements as well – notably the analogue of what are now regularly called 'pitch accents', and a feature [stress] that accounts for distinctions of vowel reduction in English unstressed syllables – but at its core lies the suggestion that the phonological

essence of stress is relative prominence in a hierarchical structure. As Liberman and Prince (1977: 333) put it, 'relative prominence is defined between phonological constituents, rather than on individual segments'.

The very plausibility of this idea makes it clear – by comparison – that nothing of the sort is true of lexical or grammatical tone. There is a clear difference of meaning in Chinese between *tāng* ('soup') and *táng* ('sugar'), but it does not in any way depend on the phonological relation between either word and anything else in a given sentence. The two words – and ultimately the two tone phonemes – are categorically distinct members of a paradigmatic set, exactly like the difference between *táng* ('sugar') and *tóng* ('brass'). The issue of whether comparison between two points in real time is involved in the perception of such distinctions – which is what was emphasized by the *JFH* distinction between 'prosodic' and 'inherent' features – is structurally irrelevant. Instead, the distinction between syntagmatic and paradigmatic suggests that Trubetzkoy's 'rhythmic-melodic' features do not form a natural class of phenomena.

3.4.6. Slower vs. faster periodicity

Finally, new evidence from a very different quarter points to the possibility that a distinction between segmental and suprasegmental may be built in to the way we process speech at the neural level. It has been proposed (e.g. Poeppel 2003, Giraud et al. 2007, Giraud and Poeppel 2012) that neural oscillations ('brain-waves') at different characteristic frequencies naturally promote attention to aspects of the speech signal with different temporal characteristics. In Giraud and Poeppel's words, there is 'a principled relation between the time scales present in speech and the time constants underlying neuronal cortical oscillations that is both a reflection of and the means by which the brain converts speech rhythms into linguistic segments' (2012: 511). Specifically, gamma waves, with a typical frequency of 25-35 Hz, would facilitate the processing of segments and some sub-segmental acoustic events such as the closure and release phases of a stop consonant; theta waves, with a typical frequency of 4-8 Hz, would focus on the global properties of syllables in sequence; and delta waves, with a characteristic frequency of 1-3 Hz, would be relevant to detecting the properties of phrases. That is, there may be a biological basis for the hierarchical arrangement of phonological domains, and for the view that certain types of phonetic properties intrinsically belong to domains of different sizes. If this is true, it provides an argument for treating 'prosody' not simply as an artefact of alphabetic writing, but as a phenomenon that can be distinguished from segmental phonology on independent grounds.

3.5. So what *is* prosody?

One justifiable conclusion from this list of possible definitions is that the current sense of 'prosody' really is incoherent. In that case, the expansion of the term among the ancient grammarians and in modern linguistics is essentially an artefact of alphabetic writing: prosody is a grab-bag of things that are hard to write with a string of symbols. Yet the phenomena in the prosodic grab-bag do seem to fall into two main clusters. One involves phonetic properties that are often thought of as running in parallel with the segmental string – elements such as pitch that are phonetically quite independent of segmental articulation, and elements that apply to stretches of speech

longer than individual segments – while the other involves hierarchical structure and syntagmatic relations. The oppositions discussed in the previous section can be separated into two groups along these lines: source vs. filter, non-verbal vs. verbal, suprasegmental vs. segmental, and slower vs. faster periodicity are all about parallel phonetic streams and long-domain properties of the segmental string, while the syntagmatic/paradigmatic distinction (and its forerunner prosodic vs. inherent) involve the hierarchical structure of the string itself.

It is tempting to suggest that these are the two sets of phenomena that began to be investigated in the 1970s under the rubrics 'autosegmental phonology' and 'metrical phonology'. To the extent that those two research traditions are theoretically coherent, then the modern sense of 'prosody' may be described as having two major facets rather than simply being a collection of unrelated topics. Unfortunately, as we saw in chapter 1 (section 1.2), autosegmental phonology has fallen on hard times, and in any case the topics of most interest to autosegmental phonology in its heyday (such as tone sandhi and vowel harmony) seem quite distinct from, say, long-domain properties of utterances like pitch range and voice quality. Arguably the only thing these phenomena share is precisely the fact that they are difficult to accommodate in a segmental idealization of speech. As for metrical phonology, it seems somewhat more promising, in the sense that there is a wide variety of work on the hierarchical organization of utterances and on what has come to be known as 'the prosodic hierarchy' (e.g. Hayes 1989, Jun 1998, Frota 2000). Much of this work flows quite directly from the research programme inaugurated by Liberman's thesis, and has moved beyond stress and intonation to consider a range of other manifestations of hierarchical structure, such as segmental sandhi (e.g. Nespor and Vogel 1986) and details of speech timing such as 'final lengthening' (e.g. Turk and Shattuck-Hufnagel 2007) and 'domain-initial strengthening' (e.g. Cho and Keating 2001). Yet some researchers apparently see 'metrical phonology' and 'prosodic phonology' as distinct 'frameworks'; with the notable exception of Mary Beckman's work (e.g. Beckman 1986, chapter 3; Pierrehumbert and Beckman 1988, chapter 6; Beckman 1996), there has been little interest in pursuing the goal of a unified theory of syntagmatic structure in phonology – of taking seriously what Beckman (1996: 19) calls 'the notion of prosody as raw organisational structure'.

In the long run there may be a basis for a single integrated theory of the broadly 'autosegmental' and the broadly 'metrical' aspects of speech, but it lies beyond the present state of our knowledge. My own recent statements about how the two areas fit together give some idea of how far we still have to go: in Ladd 2008a chapter 8, I argued that we need to incorporate some notion of hierarchical structure into our analysis of intonational phonology, yet in the same year (Himmelmann and Ladd 2008) I also set forth some ideas about why accentual phenomena really are different from other 'prosodic' properties . I agree with Beckman that a good general theory of syntagmatic phonological structure is an important goal, and that, in Pierrehumbert and Beckman's memorable phrase (1988: 160), 'we speak trees, not strings'. But at present it is hard to see how such a theory will be of much use in understanding, say, overall pitch range, or the language-specific phonetic details of intonational pitch movements.

When I was writing the first edition of *Intonational Phonology* (Ladd 1996), I tried to come up with a single term to describe the general approach to intonational structure

based on Liberman 1975, Bruce 1977, and Pierrehumbert 1980. I mailed a number of colleagues who were active in the field at the time and offered a reward to anyone who could improve on 'autosegmental-metrical', which I was using as a placeholder and which struck me as roughly accurate but hopelessly clumsy. No one thought of anything better, and the term 'autosegmental-metrical' stuck, albeit now usually in the abbreviated form 'AM'. The inability to provide a unified name seems symptomatic of the fact that the two clusters of phenomena are quite distinct. It is difficult to predict whether future research will show the way towards genuine integration, or whether it will confirm that we really are dealing with separate phenomena that should not be lumped together under a single heading 'prosody'. In the meantime, the word itself is apparently here to stay.

3.6. Appendix

3.6.1. Definitions of words corresponding to English *prosody* and *prosodic* in dictionaries in other Western European languages[7]

In general-audience print dictionaries of French, German, and Italian published as recently as the 1970s and 1980s, words corresponding to English *prosody* and *prosodic* are defined in ways related to the first definition in the original *OED* – the sense related to versification and metrics. In the Wahrig German dictionary (first published 1968, revised 1979), musical text-setting is mentioned in this connection; some Italian dictionaries also make reference to the classical sense of Greek word accents. In Spanish and Portuguese the traditional meaning seems to have been closer to the *OED*'s second meaning, relating to correct pronunciation, but the modern linguistic meaning was equally absent from the dictionaries. As for the Firthian sense, we would naturally not expect it to show up in dictionaries of other languages, as it was restricted to a small group of English scholars writing in English.

I have found only two clear indications of the modern linguistic sense in these dictionaries before the mid-1990s. One is in the Robert French dictionary, which quotes Marouzeau's technical dictionary cited in the main body of this chapter (but only under the headword *prosodique*, not *prosodie*):

> *prosodique*: … Phonét. *Caractéristques prosodiques d'une langue:* « Les éléments phoniques (dynamique, mélodique, quantitatif, etc.) qui caractérisent telle ou telle tranche de la chaîne parlée, par ex. dans le mot, la syllabe » (Marouzeau).
> [*prosodic*: Phonetics. *Prosodic characteristics of a language*: "Those phonic elements (dynamic, melodic, quantitative, etc.) that characterise a given stretch of the speech signal, e.g. in the word, the syllable." (Marouzeau).]

The other is in the massive multi-volume UTET Italian dictionary, edited by Battaglia and Squarotti; publication began in 1961, but the volume containing *prosodia* appeared in 1988. The first two definitions refer to the classical and poetic uses, but the third reads:

> Fonol. Nella linguistica moderna, l'insieme dei caratteri fonici (dinamici, melodici, quantiativi) che sono peculiari di una determinata sezione del discorso.
> [Phonol. In modern linguistics, the set of phonic properties (dynamic, melodic, quantitative) that are characteristic of a specific section of speech].

The wording strongly suggests that this definition is also based on Marouzeau.

3.6.2. Google Scholar™ search of titles containing *prosody* and *prosodic*

In April 2011 I carried out a search on Google Scholar for publications having the word *prosody* or the word *prosodic* in their title, and examined the changes over time

---

[7] In preparing this section I consulted the entries for words corresponding to *prosody* and *prosodic* in major dictionaries in French, German, Italian, Spanish, Portuguese, Swedish, and Danish. Details are given in a separate bibliographical list at the end of the book.

in the number of publications. In order to get an impression of the general background growth in the amount of published material in linguistics available to Google's search engines I also searched for titles containing a sample of other unambiguously linguistic terms (e.g. *pronominal, fricative, ergative* and *diglossia,* but not *morphology*, *segmental*, *labial* or *pragmatic,* which are used in other fields as well). No attempt was made to correct for duplicate entries, though a few clearly spurious entries were discarded. The results are shown in the graphs below. All the graphs show plots for eight time periods, beginning with all titles published in 1970 or any time earlier, then continuing in five-year intervals 1971-75, 1976-80, and so on up to 2001-05. In most cases there are more titles in the first period than in the second, but only because the first period covers everything from the beginning of Google's online records up to and including 1970. The overall increase from 1971-75 to 2001-05, on the other hand, is a genuine increase, reflecting both the growth in the volume of scholarly publication and the increasing availability of material online.
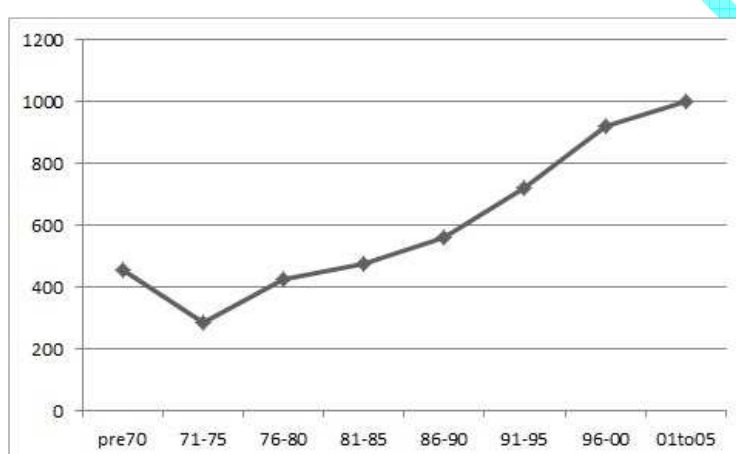

Fig. 3.1. Average Google Scholar hits for works with linguistic terms in the title. Period 1 = pre-1970; thereafter 5-year intervals (1971-75, 1976-80, etc., up to 2001-05). Terms included here are pronominal, interrogative, accentual, intonation, lexical, syntactic, phonology, fricative(s), vowel, ergative, diglossia, deixis, *and* diphthong(s).

The average increase in the number of linguistics titles between 1971-75 and 2001-05 is approximately 350% – that is, there are roughly three and a half times as many linguistics titles recorded on Google Scholar for 2001-05 as for 1971-75. The increases range from 200% for *diglossia* and *ergative* to nearly 500% for *pronominal* and *deixis*, with *lexical* an outlier at nearly 1000%. The overall graph is shown in Fig. 3.1. A different choice of search terms would obviously yield results that differ in detail, but it seems reasonable to suggest that there are three or four times as many publications in linguistics now as there were four decades ago.

By comparison to this background average growth of publication, the increase in titles containing *prosody* or *prosodic* has been much greater, in the neighbourhood of 1000%. That is, there are now ten times as many publications on prosody as there were four decades ago. This can be seen in Fig. 3.2, which shows overall uses of *prosody* and *prosodic,* without any attempt to break them down according to the sense in which the two terms are used.
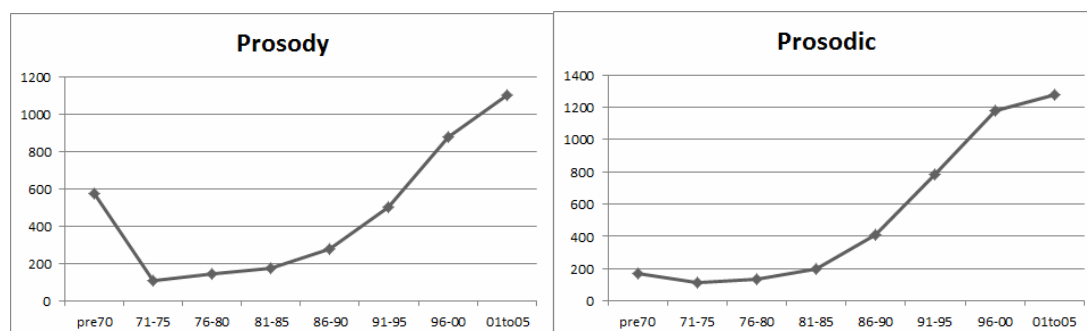
Fig. 3.2. Google Scholar hits for works with terms prosody *and* prosodic *in the title. As in Fig. 3.1, period 1 = pre-1970; thereafter 5-year intervals (1971-75, 1976-80, etc., up to 2001-05).*

To get an idea of the words' changing meanings, I did a rough two-way classification of all titles containing either word, according to whether they dealt with poetry and metrics or with senses related to linguistics or psychology, including the current linguistic sense (whatever exactly that is). The classification was done solely on the basis of the title and the brief quotes that appear on the pages of search results in
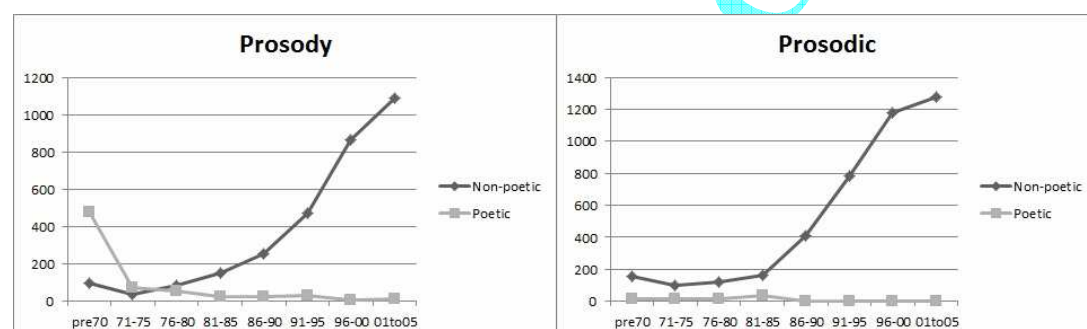

Fig. 3.3. Estimated Google Scholar hits for works with the poetic and non-poetic senses of the terms prosody *and* prosodic *in the title. As in Fig. 3.1, period 1 = pre-1970; thereafter 5-year intervals (1971-75, 1976-80, etc., up to 2001-05).*
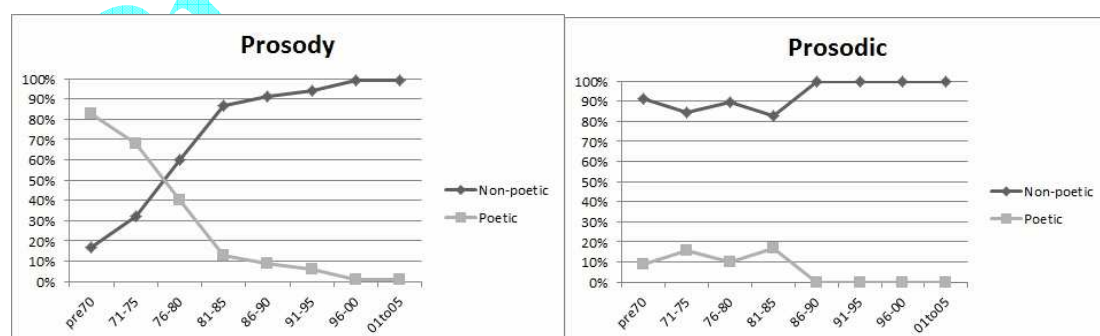

Fig 3.4. Estimated Google Scholar hits for works with the poetic and non-poetic senses of the terms prosody *and* prosodic *in the title, expressed as a percentage of the total number of hits for the two terms. As in Fig 3.1, period 1 = pre-1970; thereafter 5-year intervals (1971-75, 1976-80, etc., up to 2001-05).*

Google Scholar, and accordingly may contain some incorrect classifications; again, no attempt was made to correct for duplicate entries. Given this methodology, the numbers should be taken only as estimates, but they give a fair picture of the evolving meanings of these two terms. Figs. 3.3 and 3.4 show that *prosody* underwent a clear shift from poetic to non-poetic meanings in the 1970s, whereas *prosodic* seems never to have been used much in the poetic sense, at least in titles. For both forms it can be seen that the 1000% growth shown in Fig. 3.2 is entirely due to the non-poetic meanings. Fig. 3.3 gives absolute numbers of publications, while Fig. 3.4 shows the same data expressed as a proportion of the total number of publications.

Finally, I also attempted a more detailed breakdown of the same data, classifying the non-poetic uses as Firthian or non-Firthian, and very roughly classifying the non-Firthian uses as belonging either to linguistics (including speech technology) or to psychiatry and clinical psychology. The Firthian meaning of both the noun and the adjective completely disappears from titles after 1985. The classification of the non-Firthian meanings is too approximate to take very seriously, but it suggests that they show roughly the same rate of increase over the period in question. There continues to be a difference between the noun and the adjective, however: psychological topics account for roughly a third to a half of all the non-Firthian uses of *prosody* in titles, but only 10-20% of the non-Firthian uses of *prosodic*.