Cross-language similarities and differences in spontaneous speech patterns

William Barry
Bistra Andreeva
Institute of Phonetics, University of the Saarland
barry@coli.uni-sb.de, andreeva@coli.uni-sb.de

Quasi-spontaneous dialogues from six languages which, according to recent discussion of rhythmic types, belong to three rhythmic groups – Russian and Bulgarian as 'stress-timed', Italian and Greek as 'syllable-timed' and Polish and Czech as an intermediate 'mixed' type – were examined for the following segmental reduction phenomena: reduction of consonant clusters, weakening of consonant articulation, residual properties from elided consonants in the original context segments, phonetic schwa-isation and syllable elision. The hypothesis tested was that there are comparable reduction phenomena in all languages since all languages allow for variation in the time and effort invested in any given part of an utterance as a means to support the relative weight of elements within the information structure. This hypothesis was borne out in principle, though there were a small number of exceptions across the six languages to the occurrence of reduction types examined.

1 Introduction

The spontaneous speech patterns considered in this paper are systematic changes in the segmental structure of words relative to their 'canonical' form, i.e. the phonetic form specified in the lexicon. Our basic hypothesis is the universality of phonetic reduction, which reflects the 'tug-of-war' between the need to convey the desired message and principle of articulatory economy (Kohler 1979, 1990, Lindblom 1990). In Lindblom's (1990) terminology, these two forces governing our speech production are the OUTPUT CONSTRAINTS of the communication partners and situation, and the SYSTEM CONSTRAINTS of our articulatory system. Efficient communication represents a state of balance between the two forces both from the situation-bound, medium-term viewpoint (speakers adapt their overall speaking style to the situation) and short-term, within an utterance. Within the bounds of the present study, we concentrate on LOCAL hypo-speech events within an utterance, namely reduced forms. Information weighting within an utterance presents continuously changing output constraints which place varying demands on the production system, calling both for more prominent, and allowing reduced, non-prominent forms along the 'hyper-hypo continuum'.

The simple corollary of our hypothesis is the occurrence of comparable reduction

DOI:10.1017/S0025100301001050 Journal of the International Phonetic Association (2001) 31/1

phenomena in all languages, since all languages allow for variation in the time and effort invested in any given part of an utterance, and these are assumed to be the defining factors for 'hypo-speech', i.e. for the production of reduced forms of words.

Given comparable production conditions – though, admittedly, these are ultimately uncontrollable since the individual reaction of each speaker to the speech task is part of the equation – any observed differences between languages should be interpretable in terms of differences in language type. Since reduction phenomena are associated with reduced time and articulatory effort, and these aspects of production are correlates of de-stressing and de-accenting, the rhythmic type to which a language belongs is usually taken to be a major factor behind its propensity for reduction. In order to assess this assumption, and before looking at the speech data we have analysed, we consider what specific implications the rhythmic organisation of a language can have for reduction phenomena.

1.1 Rhythmic types

Traditional rhythmic typology distinguishes stress-timed, syllable-timed and mora-timed languages (Pike 1945). Stress-timing is seen as being strongly conducive to reduction phenomena (Gimson 1989[4]: 263ff., Kohler 1995[2]: 117) because the prominence given to the stressed syllable and the ostensibly equal duration of the stress groups lead to a temporal compression of the unstressed syllables between them. Conversely, the theoretical equality of the syllables in a syllable-timed language offers equal resistance to reduction for all the syllables, though clearly lexical schwa-syllables, when they exist, e.g. in French *le petit* (/ləpəti/ → [ləpti]), are clearly an exception to the equality principle, as the schwa-elision rules for French document.

The impressionistically based di- or trichotomy of prosodic types has been subject to many attempts to find a measurable basis. These were initially orientated towards the syllable- or stress-isochrony postulated by Abercrombie (1967) as the strongest form of Lloyd-James' original 'morse-code' vs. 'machine-gun' distinction (Lloyd-James 1940, reported in Pike 1945: 17). However, the instrumental investigation of isochrony (e.g. Uldall 1972; Pointon 1980; Wenk & Wioland 1982; Dauer 1983, 1987; Hoequist 1983a,b) has never provided strong support for the claims. Interstress intervals in English have been shown to be proportional to the number of syllables in them (e.g. Bolinger 1965) and even using different, theoretically probing measures, syllable duration is not constant in syllable-timed languages (cf. Hoequist 1983c for Spanish). Currently, stemming from work by Dasher & Bolinger (1982) and Dauer (1983, 1987) there is growing acceptance of the view that the rhythmic character of a language is the product of a number of specific segmental phonological properties. This has led to a relaxation of the strict trichotomy of prosodic types, at least for the syllable- vs. stress-timed distinction. The observation (e.g. Nespor 1990) that vowel reduction + complex syllable structure (associated with stress-timing) and non-reduction + simple syllable structure (associated with syllable timing) are not the only possible combinations, even in European languages, necessarily changes the view of rhythm typology. Languages like Polish, which has very complex syllable structures but no vowel reduction, and Catalan and Portuguese, which have the relatively simple syllable structures of the Romance languages yet have a clear system of vowel reduction, cut across the normal grouping of segmental properties associated with syllable- and stress-timing. Also, measuring the proportion of time spent on vocalic and consonantal segments rather than searching for isochrony (Ramus et al. 1999, Ramus & Mehler 1999, Grabe & Lowe to appear) has demonstrated that there is indeed a relatively simple instrumental basis for distinguishing not only the traditional rhythmic language types, but also (some of) the languages with mixed structural features. Finally, perception experiments have shown that adults can distinguish languages on the basis of resynthesized [sa], con-

sonant-vowel duration patterns (Ramus et al. 1999, Ramus & Mehler 1999), and evidence is beginning to accrue that newborns can recognise these same vowel-consonant differences between rhythmic groups (cf. the discussion in Ramus et al. 1999).

1.2 Rhythmic types and reduction

While the new view of rhythmic types and the instrumental measures supporting it have broken up the traditional categories, they have also confirmed the existence and defined groups of languages that can be associated with the traditional 'stress-timed' and 'syllable-timed' rhythmic types. But since it is more convincing to regard rhythmic type as the product of a number of phonological properties, not as a phonological prime, as the traditional tripartite division is often understood, the question in the context of this study is: Which structural properties associated with rhythmic types might affect reductions in continuous speech, particularly in spontaneous speech?

Consonantal complexity, in terms of onset and coda clusters, is an obvious factor in reduction processes. It gives a wider range of syllable types and more inherently heavy syllables, and is therefore naturally associated with stress-timed languages. Importantly in this context, the more complex a cluster, the more room for articulatory saving, assuming word identity is not endangered. At the same time, under comparable prominence conditions, a more complex syllable will result in a phonetic shortening of the vowel with concomitant spectral change, a tendency to centralise.

Lexicalised vowel reduction, on the other hand, or stress-dependent 'vowel-quality alternation', as we prefer to call it, is a property which is also traditionally associated with stress-timed languages and appears to be implicitly linked to non-phonological reduction processes in continuous speech. But in fact, whatever its origin historically, it is a lexically fossilized process and offers no scope for phonetic reduction other than being a potential site for deletion (cf. Farnetani & Busà 1999). It is found both in languages that are considered stress-timed and in syllable-timed languages, and conversely there are stress- and syllable-timed languages that do not have vowel-quality alternation. Reduction in this sense is best documented in the alternation of vowel categories within the same morpheme or morpheme sequence, under conditions of destressing. A peripheral vowel can alternate with a more central, often schwa-like vowel or a non-close front or back vowel can alternate with a closer vowel. For example, there is an alternation between /ɒ/ and /ə/ as a function of lexical stress in the English words *philosophy* /fɪˈlɒsəfi/ and *philosophical* /fɪləˈsɒfɪkəl/. German, on the other hand, pronounces the <o> in both stressed and unstressed syllables in *philosophisch* (stressed on <soph>) as [o], the two vowel sounds differing in length, but – in auditory impressionistic terms – not quality, as a function of stress. Russian /o/ is realised as [a] in unstressed position; Bulgarian and Portuguese unstressed /e/ and /o/ are realised as [i] and [u], respectively. Bulgarian, Catalan, English, Portuguese and Russian are examples of languages with vowel-quality alternation, whereas French, German, Italian, Polish and Spanish do not have it. The occurrence of continuous-speech reductions in a language without vowel-quality alternation, namely German, many of them comparable to reductions in English, which does have vowel-quality alternation, are abundantly documented (Kohler 1979, 1990; AIPUK 30, 31, 32 (1996–97). The status as a traditionally stress-timed language appears more important as a basis for phonetic reduction in continuous speech than the status as a language with vowel-quality alternation.

Should we therefore expect traditional syllable-timed languages to resist hypo-speech reductions? For a number of reasons, the answer is no. Since continuous-speech reductions are attributed to reduced articulatory effort, and the status of a language as syllable-timed in no way precludes the accenting of words within phrases as part of the information structuring process, there will necessarily be non-accented words offering

potential for reduction. The choice of phonetic means to achieving the necessary prominence is the same as for stress-timed languages, namely tone, duration and intensity. It is therefore likely that shorter and less intense unaccented syllables are also produced with less articulatory effort. Thus there is clearly potential for differences in the articulatory effort invested in different syllables of a phrase, whether the language be traditionally syllable- or stress-timed.

Not all syllables between the accent-bearing syllables of prominent words are equally open to reduction, however. There is a further level of (potential) prominence, in that polysyllabic words can still have a lexically stressed syllable which may retain some degree of prominence when unaccented. This appears to be generally true for stress-timed languages, but it is also true for syllable-timed languages such as Italian, Spanish and Greek (though not French) which – in implicit contradiction to their supposed syllable-timed status – have a word-defining stress pattern. This offers at least the theoretical possibility in both stress-timed and syllable-timed languages of differentiating between the unstressed and stressed syllables of non-prominent words as more or less likely sites for reduction.

Finally, if the language is also a language with vowel-quality alternation, it is theoretically possible to distinguish further between unstressed full vowels and unstressed reduced vowels, since not all unstressed vowels are necessarily maximally reduced. The effect of this difference on continuous speech reductions can only lie in the potential for extreme shortening and ultimate elision of the reduced vowel.

In summary then, no single rhythmic type, whether traditionally stress- or syllable-timed or of a mixed type, would appear resistant to phonetic reduction processes. All utterances with unaccented words offer potential for vowel and consonant reduction, though languages with a stressed-unstressed distinction at lexical level should be more prone to reduction in the unstressed syllables.

Specifically, we expect shortening and centralisation, ultimately 'schwa-isation' of vowels, loss of consonants from clusters, and general weakening of consonant articulation. Consonant reduction in clusters is a de facto reduction of syllable complexity. Therefore in languages with greater syllable complexity the potential for economy of effort is greater. However, wherever a coda and a consonantal onset form a cluster there is potential for reduction if it can be argued that articulatory effort is thereby reduced. So in general we may assume differences between languages in degree and frequency of reduction but not differences of occurrence vs. total non-occurrence (except in the limiting case that a strict CV-language would provide no structural basis for e.g. cluster reduction). Reduced effort can, however, also be reflected in changes to single-consonant realisation. Lenition of fortis consonants, fricative realisation of stops and approximant realisation of fricatives all represent a weakening that can apply to individual consonants, and are reduction processes which are not restricted to languages with complex syllable structures.

1.3 Non-phonological factors

A feature of those languages whose reduction patterns have been extensively studied, namely English and German, is their richness in isolatable monosyllabic grammatical elements such as articles, pronouns, auxiliary and modal verb forms and their abundant prepositional systems. Since these elements are accentable but unaccented by default they have in fact provided the focus for the analyses of reductions, and have served to illustrate the relationship between redundancy of information, deaccenting and formal reduction. In languages where some of these elements are less frequent or non-existent, the frequency of reductions, though phonetically similar when they occur, will inevitably be much lower. However, no significant differences stemming from radical structural differences of this kind are to be expected in the languages examined here.

1.4 Spontaneous speech

So far, the potential for reduction has been considered in 'continuous speech', independent of the type of speech produced. 'Spontaneous' speech is difficult to define in a watertight manner, but it should be possible to accept that it is unscripted, unprepared in terms of the number, organisation and expression of the information points it communicates. The speaker decides ad hoc what he or she wants to communicate at each stage in the communication situation, plans and utters it. The stage may be defined by the speaker's reaction to a non-speech event, to what he or she has just said (e.g. a correction or continuation), or to an utterance from someone else (e.g. a turn in a dialogue). It is therefore different by definition from a read text, a prepared but unscripted talk, or an acted dialogue, though informal listening tests that we have carried out have shown that a 'good' reader, speaker or group of actors can produce non-spontaneous speech which is difficult to identify as such. Can we, therefore, expect to find different types of reduction in spontaneous speech?

Certainly if spontaneous speech is transcribed and then read by the same speakers, there are differences at many points in the text (quite apart from the impossibility of reproducing any dysfluencies that may have occurred in the spontaneous text). However, we have so far not identified reduced forms that are peculiar to one particular form of continuous speech. In reading, it is immediately noticeable that few readers (with the exception of professional, dramatic readers) vary their tempo to any great extent. In spontaneous speech, on the other hand, tempo changes, particularly in dialogues, are extremely common. This is directly explainable in terms of the relative importance of the information chunks to the speaker. A written text exists as a linear sequence of words. Most readers, even fluent ones, only decode the primary meaning and therefore differentiate prosodically the content and function words, and even the focus of each successive clause. But in spontaneous speech the logical links and consequent weighting of one information chunk relative to another – which most readers of a text do not decode and express – is the point of departure for the planning and execution.

The direct hypothesis from this is that spontaneous speech will contain a wider range of forms, from very reduced to non-reduced hyper-forms, whereas less spontaneous text types will be more uniform.

2 Speech material and analysis

2.1 Languages

The languages examined in this study were selected on the basis of coverage of the properties discussed in the introduction and of accessibility to speaker pairs for dialogue recordings under comparable conditions. The languages and their defining properties are summarised in table 1 below. Four of them, the Slavic group, are usually considered to be stress-timed, though under the new classification criteria, Polish and possibly also Czech may be of the mixed type. The other two languages, Italian and Greek, are usually considered to be syllable-timed. In terms of lexicalised vowel reduction, the two non-Slavic languages do not have vowel-quality alternation in accordance with the traditional syllable-timing assumption, Russian and Bulgarian conform to the stress-timing expectation, i.e. they DO have vowel-quality alternation while Polish and Czech do not.

In terms of syllable complexity, all four Slavic languages allow complex syllables, both in onsets and codas, while the two non-Slavic languages allow only a small number of onset clusters and only single-consonant codas. These properties clearly group the six languages into three structurally differentiable pairs.

Table 1 Languages used in the study.

Russian:	E.	Slav	vowel alternation	complex syll. structure
Bulgarian:	S.			
Polish:	W.			
Czech:	W.			
Italian:	Romance		no vowel alternation	simple syll. structure
Greek:	Independ.			

2.2 Speakers and speech material

Recordings were made of at least two speakers per language. Speaker selection was based on availability. The material recorded consisted of quasi-spontaneous dialogues (staged 'map tasks' (Anderson et al. 1991) or comparable verbally interactive spatial tasks) together with the read versions of the transcribed dialogues and the words of dialogues read in citation form. The dialogues lasted anything from 2 to 5 minutes. Each person read part or all of their own half of the dialogue (depending on its length) and subsequently the list of words contained in that part of the dialogue.

2.3 Analysis of reduction phenomena

The dialogues were transcribed orthographically and the words rendered phonemically according to their canonical form. Phonetic realisations were transcribed auditorily supported by visual inspection of the microphone signal and the spectrogram. In accordance with the understanding of phonetic reduction discussed above, tokens which deviated in one or more of the following ways from the canonical form were identified:

a) simplification of consonant clusters,
b) weakening or loss of single consonants,
c) spread of residual features from 'elided' segments,
d) centralisation of vowels ('schwa-isation'), and
e) temporal reduction and qualitative fusion of vowel sequences.

A further category is presented and discussed which does not immediately fall under the same heading, namely,

f) intonation-dependent reduction.

3 Results

In presenting and discussing our observations of the data, we are unable to present spectrogram examples of each type for each language in the available space. Therefore we select for each type of reduction, give examples from structurally different languages and comment if certain types were not found in certain languages. (See http://www.coli.uni-sb.de/~wbarry for the audio recordings and spectrograms of the example utterances.)

3.1 Simplification of consonant clusters

From the articulatory side (input constraints) there is a clear 'economy of effort' effect in reducing the number of consonants produced intervocalically. In terms of probability of reduction, the greater degree to which a sequence of gestures is established WITHIN a word compared to BETWEEN words would suggest a greater probability of reduction in cross-word-boundary clusters. Taking perceptual factors into consideration (output constraints), there is a clear onset vs. offset difference in information loss if consonants are dropped. Thus cross-word-boundary clusters would be expected to lose coda-consonants rather than onset consonants. Cutting across this gradation of phonetic information is, however, the contextual predictability of a particular syllable, either as a monosyllable in the syntacto-semantic sequence or as a syllable in a multisyllabic word. Our data showed cluster reduction and consonant weakening within clusters to be common across all six languages. The examples in (1)–(6) illustrate.

(1) Polish: /jak ˈjɛstɛʃ/ ('as if you were') → [jăˈkĭsə̯ʃ]
(2) Bulgarian: /ˈdɔlnata tʃast prɔsto/ ('the bottom part simply') → [tʃaːspɾɔsʊ]
(3) Czech: /ve ˈvzdaːlenɔsti/ ('at a distance of') → [fvɛˈzdaːlɔsi]
(4) Italian: /riˈspɛtːo al denˈtista/ ('relative to the dentist's') → [denˈtisa]
(5) Greek: /aˈnamesa stus ðiɔ ˈdrɔmʊs/ ('between the two roads') →[sɵʑjð] (see figure 1)

The loss of /t/ from /st/ clusters was as prevalent in the 'syllable-timed' as in the 'stress-timed' languages, intervocalically as well as in /stC/ clusters. The loss of /t/ in intervalic clusters means, in effect, that the onset of the second cluster is changed. This was generally an unstressed syllable of a multisyllabic word, but in the above Greek example, a word-onset cluster in /stus/ is reduced to /s/. Perhaps coincidentally, no reduction of /st/ clusters were found in Russian except for the frequent form [tɔɪs] for /to estʲ/ ('that is'), though other clusters, in particular multi-consonant clusters with /v/ and /f/ were prone to reduction.

(6) Russian: /na ˈurovnje ˈbelovo ˈdoma / ('at the level of the white house') → [naˈuərnʲi]

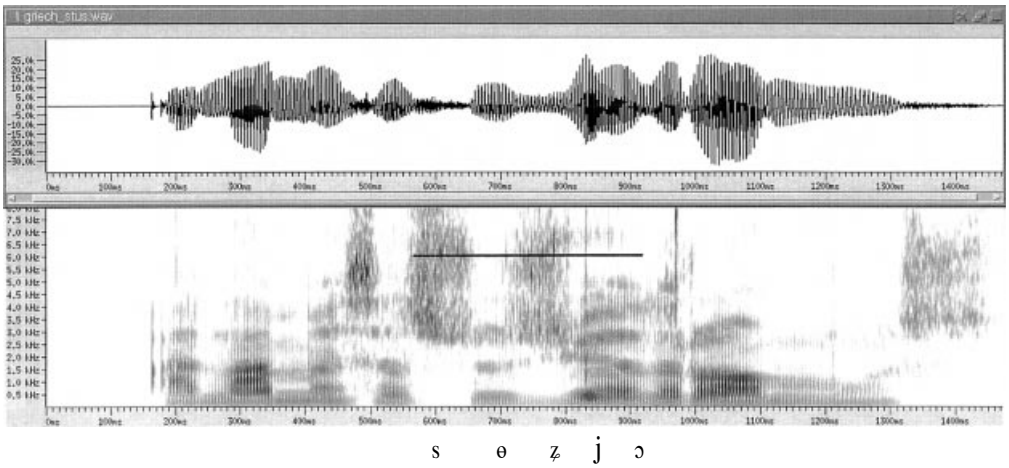

Figure 1 Cluster reduction in Greek.

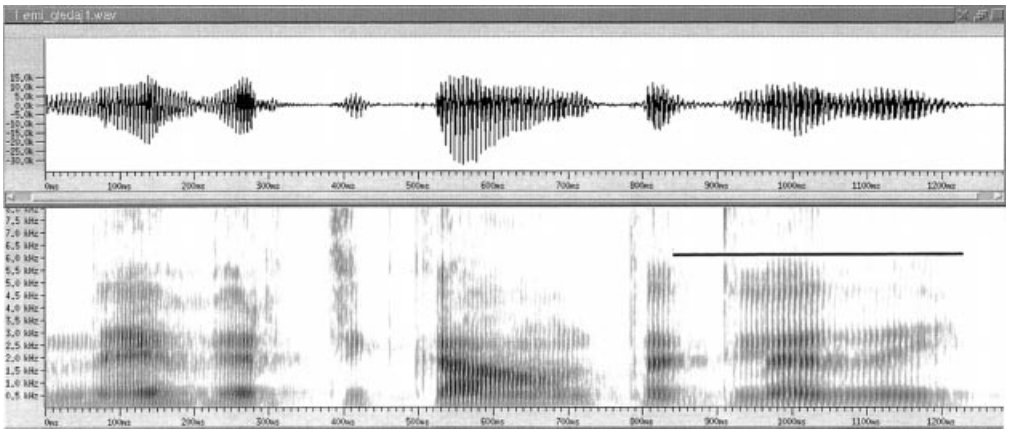

Figure 2 Alveolar weakening in Bulgarian.

3.2 Weakening or loss of single consonants.

The lenition of fortis stops, the fricativization of lenis stops and the corresponding weakening of fortis and lenis fricatives to lenis fricatives and approximants, respectively, mark stages of consonant reduction that can be captured in transcription. Alveolar consonants are known to be particularly prone to reduction and this was borne out in the data. The previous section showed that, across the languages, it was often /t/ which was elided in clusters. /t/ singletons were not elided in these recordings but /d/, on the other hand, as the lenis member of the opposition, was prone to reduction, either reducing to a short tap, a weak fricative or approximant, or disappearing completely. In five out of the six languages there were numerous cases of weakened or deleted /d/, but none were observed in Greek. The following examples may serve as illustration:

(7) Bulgarian: /ˈnɛ vɛrtiˈkalnata ˈglɛdaj/ (‘dont’s look at the vertical’) → [ˈɣlɛᵈaj] (see figure 2)
(8) Russian: /nat ˈurovnje/ (‘above the level’) → [nãˈd̪ŭrɔ̃nʲʊ]
(9) Polish: /nɔ ˈdɔbʒɛ / (‘all right’) →[nəˈd̪ɔbʑə]
(10) Czech: /ˈdveste ˈsedumdesaːt ˈstupnjuː (‘two hundred and seventy degrees’) →[sɛ̯ɾŭməsd̪ːt]
(11) Italian: /al di ˈsopra dei bamˈbini/ (‘above the children’) →[ˈsɒpɾɛ̯i]

(For reduction of /bamˈbini /, see section 3.3 below.)

3.3 Residual and spreading features from weakened or elided segments

If the primary lingual articulation is elided, any secondary (or otherwise accompanying) articulation can remain as a residual signal of the intended sound to colour the original context segments. This is particularly prevalent with laterals. The vocoid realisation of /l/ is common, post-vocalically (see: S.E. Engl. *milk* [mɪʊk] or *tickle* [ˈtɪkʊ]; German *Milch* [mɪjç] or *Million* [mɪˈjːoːn]; and reflexes in French of old ‘dark /l/’: *chateau*, etc.). Examples of /l/ elision or vowelisation were found in four of the five languages where the reduction is possible (Polish cannot be included because of the phonologised use of

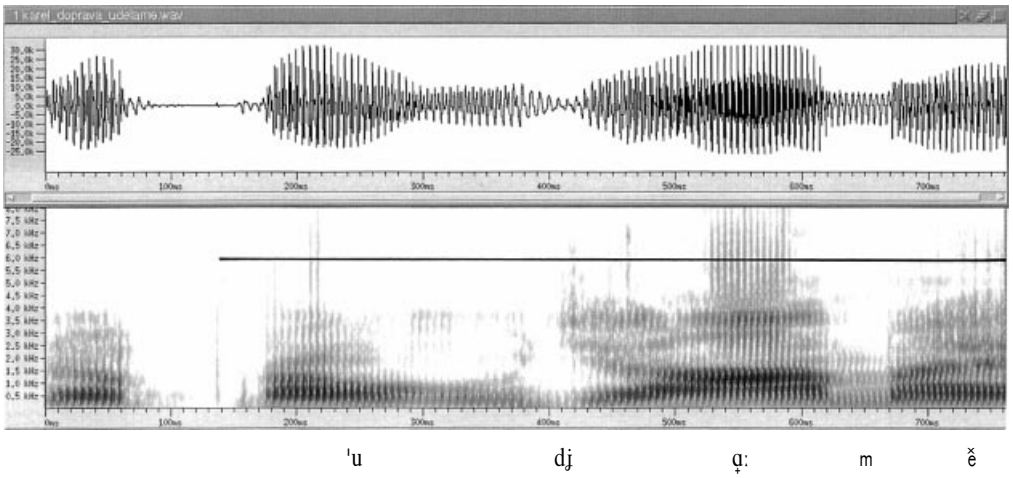

Figure 3 /l/ elision in Czech.

[w] for historical dark /l/). No evidence of vocoid /l/ was found in Greek, where /l/ does not occur in the coda.

In Italian, the only examples of voicoid /l/ realisation (or elision) were following a close palatal vowel, i.e. in the article *il* before a consonantal onset.

(12) /tra il ˈfiume e la ˈstrada/ (‘between the river and the road’) →[traˌiˈfːʲuːmɛ̆]

In Czech too, elision of the lateral was only observed after a palatal:

(13) /ˈudʲelaːme/ (‘we make’) → [ˈudʲɑ̟ːmɛ̆] (see figure 3)

In these cases, as in fact with German, the vocoid /l/ variant appears to be facilitated by the existence of surrounding palatal segments.

In Russian and Bulgarian, on the other hand, the dark /l/ was frequently produced as a vocoid, irrespective of the palatal or velar quality of the preceding vowel, with a considerable retraction of the vowel if it was a phonemically front vowel. This has parallels to the vocoid /l/ in English, where there is some evidence that it is more likely to occur following palatal vowels as in *milk* (cf. Hardcastle & Barry 1989), where the velarity provides a greater perceptual contrast. However, there is also the possibility of velar harmony through the syllable: [mθ̟ʊk]. In the following Russian example, the /o/ preceding the canonical /l/, which should be realised as [ɐ], takes over the velarity of the vocoid /l/ while the following /u/ is strongly centralised.

(14) Russian: /u mjenja polutʃilosj tak/ (‘for me it came out like that’) → [pʊ̆ɵˈtʃəʊ̯ɐ̟ʐ̥]
(15) Bulgarian: /ˈɔkolo ɛˈdin poloˈvina/ (‘approx. one and a half’) → [ˈɔkʊ̆ːɵˈdʲĭnipʊ̆ːˈʊinɐ̟ːː]

The strong tendency for vocoid /l/ realisation in Bulgarian, even in stressed and accented syllables, might suggest a phonologisation similar to that found in Polish except that the speaker of the words in (15) alternated in another utterance between a lateral and a vocoid in the same word /kəˈtʃɪlə/ (‘climbed’) producing a lateral in the SPONTANEOUS utterance and a vocoid in the READ version of the same utterance.

(16) /kɐˈtʃilɐ/ (‘climbed’) → [kəˈtʃəᵚə] vs. [kəˈt̪ʃɪlʸə] (see figure 4)

Other cases of residual reflexes of elided segments found in the data were nasality and palatality. The Italian example (11) in the previous section illustrates the

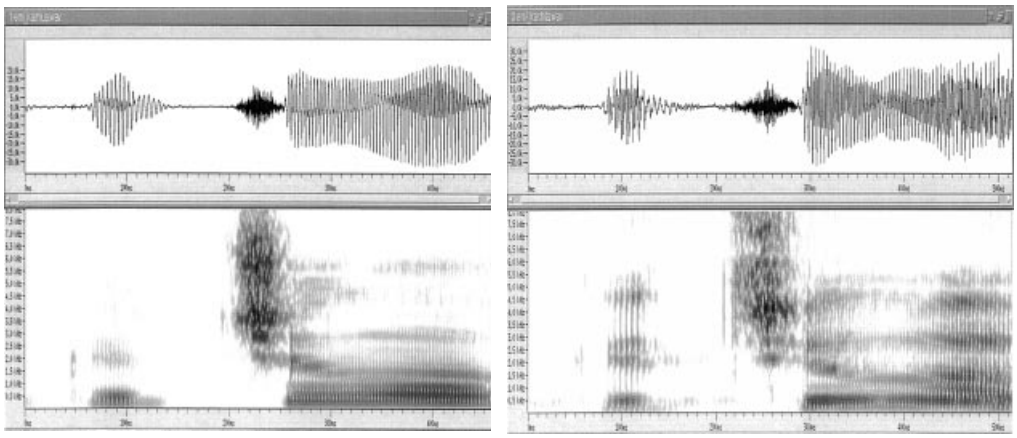

Figure 4 Vocoid and lateral realisation of /l/ in Bulgarian.

phenomenon of /n/ elision with a spread of nasality to neighbouring vowels. What makes the Italian example of /n/ elision in *bambini* [bambĩːi] noteworthy compared to the usual example of *can't* [kæ̃t] given for American English nasal spreading is its intervocalic position, the two-way spread of nasality. In Polish, the loss of a palatal approximant /j/ in /gdʑe jɛst/ was accompanied by a spread of close palatality to the remaining vocalic nuclei.

(17) /a ˈgdʑɛ jɛst ˈbomba/ ('where is the bomb?') → [gd̥ʲɨ̆ːz]

In both these cases two phonological syllables are fused in one sonority peak at the phonetic level, though of course the greater length of the vowel that remains makes the decoding of the underlying structure unproblematical in context.

3.4 Schwa-isation

Examples of centrally reduced ('schwa-ised') vowels were found in the languages WITHOUT as well as in those WITH vowel-quality alternation where other reduction targets were expected.

In Russian, unstressed /a/ and /ɔ/ are said to merge in /ʌ/, and high vowels are slightly lowered – /i/ to [ɪ] and /u/ to [ʊ] (Timberlake 1993). We in fact observed a strong, and relatively undifferentiated, centralisation tendency, which the following phrase serves to illustrate:

(18) /pɔ vʲerti ˈkali/ ('on the vertical') → [pə̯vərtɨˈkalə̥]

In Bulgarian, there is no general agreement in the literature as to the standard vowel alternations with destressing (compare Ternes & Vladimirova-Buhtz 1999; Tilkov & Bojadzhiev 1977; Scatton 1975, 1993). A centralising change of /a/ and /ɤ/ to [ɐ] is accepted in all three descriptions, but Tilkov claims schwa as a second-level reduction. Front and back, mid and close vowels /ɛ/ and /i/, and /ɔ/ and /u/ are said by Scatton to merge when unstressed, though he gives no phonetic value. Tilkov and Ternes & Vladimirova-Buhtz agree that the back mid vowel /ɔ/ is raised to [o], but Tilkov again notes a shift to [u] at the second level of reduction, whereas Ternes & Vladimirova-Buhtz state that /ɔ/ and /u/ merge in [o] when unstressed. Although the

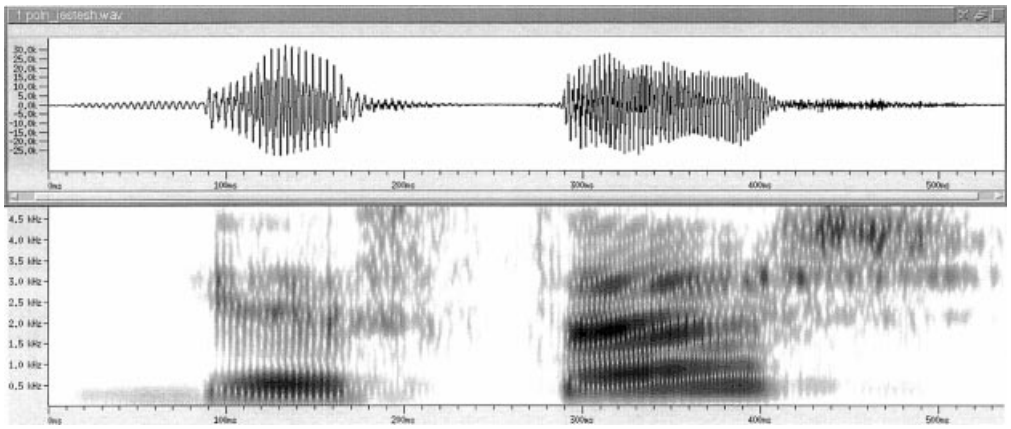

Figure 5 Schwa-isation in Polish.

isolated-word productions confirmed the [ɐ] weakening and the raising of the back mid vowel, the spontaneous utterances showed many schwa-like realisations irrespective of the underlying vowel, as, for example, in (19).

(19) /kɐ'tɔ si sɛ kɐ'tʃiɐ/ ('when you have climbed') → [kətəsːɐ̟kə'tʃəlˠə]

In the four languages without vowel-quality alternations, whether the traditionally syllable-timed Italian and Greek, or in Polish and Czech, there were also clear cases of schwa-ised vowels:

(20) Czech: /tak ten tsil 'bude nad/ ('so this ''finish'' is over?') →[tə̟n]
(21) Italian: /il 'fiume lo 'devi la'ʃare/ ('you must leave the river') → [i'fjumə̟lə̟re'vilă'ʃără]
(22) Greek: /akri'vɔs ðipla apɔ tɔ/ ('right next to') → [ɐɣəɾə'vɔz̟ɪplɐptð]
(23) Polish: /'jesteʃ/ ('are you?') → ['jɐ̟stə̟ʃ] (see figure 5)

The phonetic nature of these reductions is shown by the residual differences in the strongly centralised vowels weakly reflecting the area of the vowel space in which the underlying vowels are defined.

3.5 Temporal reduction and qualitative fusion of vowel sequences

Except for languages which mark vocalic word onsets with a glottal stop (often reduced to a phase of glottal constriction), there are frequent vowel elisions and consequent syllable loss when a word-final vowel abutts a word-initial vowel. Although the frequency of cross word-boundary vowel sequences varies considerably among the six languages examined here, being clearly more frequent in Greek and Italian, all of them presented examples either a) of elision of one vowel, or b) of a fused vowel quality ascribable to a mixture of the two vowels contained in the canonical forms. Czech differed from the other five languages in its tendency, like German, to mark vocalic word onsets glottally, though there were still many cases of fusion of unaccented function word elements with the vocalic onsets of accented lexical units. The Russian phrase,

(24) /kak bɨ u mje'nja/ ('somehow with me') → [kagbə̆ŭmĩ'jə̆]

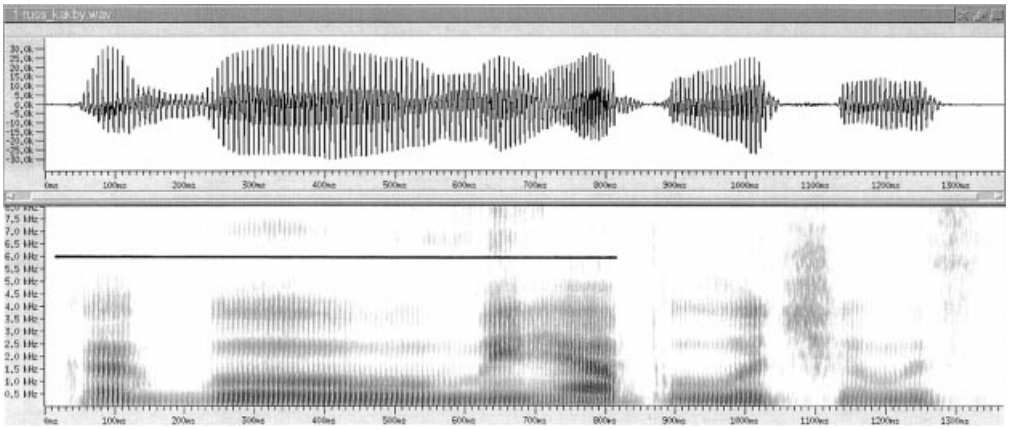

Figure 6 Mutual modification of abutting vowels in Russian.

shows a case of mutual colouring of the two vowels without complete loss of either, although both are very short and, as might be expected, also centralised (see figure 6).

In the Polish

(25) /nɔ jak tɔ̲ alɛ 'jɛstɛm 'tɛraz vɨ'sɔkɔ/ ('but I *am* at the top!') → [ta̲'lɛ̰sᵗəm]

there is a loss of the back mid rounded /ɔ/ without any noticeable rounding of the /a/, whereas the Bulgarian

(26) /'tri santi'metra̲ ɔt'gɔre/ ('three centimetres above') → ['metɾ̆ɒ̆d]

shows the reverse effect, and even the raised quality expected of unstressed /ɔ/, though also some centralisation. In the Czech

(27) /kdiʃ si̲ 'u̲dʲelam . . ./ ('when I make . . .') → ['s̪ʉ̟dʲɑːm]

we observe a temporal reduction and qualitative fusion of the two underlying vowels, while in Greek

(28) /i̲nɛ ɛna̲ kɔ'ritsi/ ('is a girl') → [in̪ənɐ]

the two vowels are reduced to one short raised schwa. Finally, in Italian

(29) /pi̯u i̲n 'alto/ ('higher up') → [pi̯n'alto]

the /u/ is elided completely, leaving a very short [i] for the underlying /i̯ui/ sequence.

3.6 Loss of syllables

All six languages had numerous examples of elided syllables. Five of the previous six examples show syllable loss across the word boundary due to vowel merger or elision, while examples (11), (13), (17) and (19) above illustrate syllable loss from other causes which accompany the reductions being discussed. An extreme example in Italian is a word-final /ne/ followed by the function word /dela/ which reduces to [nela]:

(30) /in diretsione dela/ ('in the direction of') → [ʔindiɾetsə̪ə̪nelɐ] (see figure 7)

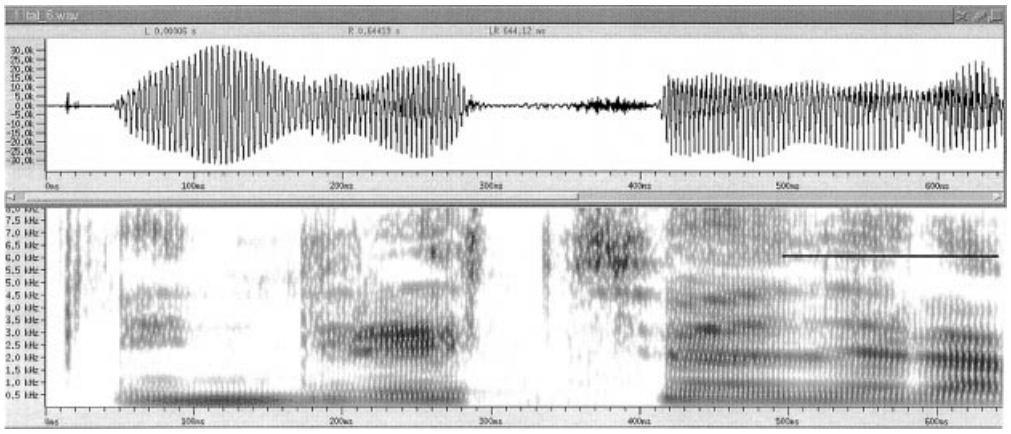

Figure 7 Syllable elision in Italian.

3.7 Intonation-dependent reduction

Russian and Bulgarian showed a propensity for syllable loss or massive reduction in phrase-final FALLING contours, while the same words in (presumably 'marked') phrase-final RISING contours retained a near canonical form, as (31) and (32) illustrate.

(31) Russian: /ˈurovnʲe/ ('level/height') → [↗ˈŭrəvᵐnʲə̯] but [↘ˈʊrəʊn] (see figure 8)
(32) Bulgarian: /na ˈdʲasno/ ('to the right') → [nɜ ↗ˈdʲasno] but [nɜ↘ˈdʲasⁿ]

4 Discussion

The initial hypothesis that the same sort of reduction phenomena will be found in fluent speech (here we considered specifically one form of spontaneous dialogue) across

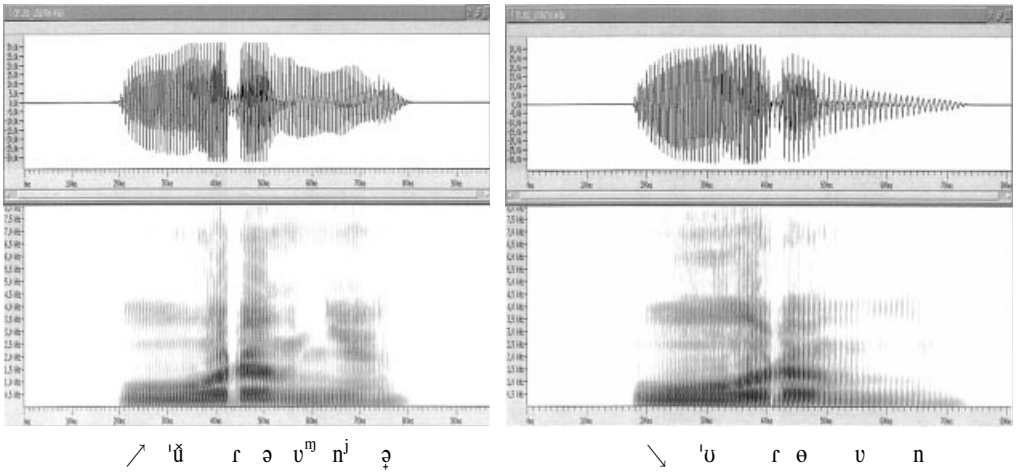

Figure 8 Intonation-dependent syllable elision in Russian.

different language types was borne out in principle. There were very few exceptions across the six languages to the occurrence of reduction types examined.

The loss of utterance-final syllables in the context of a final intonation fall, a phenomenon not previously mentioned in the literature, was only found in Russian and Bulgarian in the material examined here. Since there is a plausible connection between an utterance-final fall and a decrease in overall articulatory effort, it remains to be seen whether the phenomenon is a feature that has developed to a more extreme form in these two languages than is the natural consequence of a 'closing' gesture.

Considerable variation in the phonetic result of /l/-elision and vocoid realisation of /l/ was also found. Here again, Russian and Bulgarian manifested a strongly established vocoid variant of their velarised /l/, almost to the exclusion of a lateral variant except under hyperspeech conditions (isolated words or strongly emphasised words in spontaneous speech). In contrast, the clear /l/ languages only vowelised the lateral in palatal contexts in extreme deaccented contexts.

Consonant-cluster reduction is clearly more frequent in cluster-rich languages, but the all-pervading occurrence of e.g. /st/ reductions, and the ubiquitous signs of consonantal weakening support the notion of universal loss of syllabic definition under the reduced articulatory effort of deaccentuation. This is further supported by the general occurrence of 'schwa-isation', independent of the rhythmic type of the language. Vowel fusion across word boundaries was also found in all the languages, though it was less frequent in Czech, even than in the other coda-rich languages, due to the frequent realisation of the glottal word-boundary marker in the stressed initial syllables with vowel onset.

Differences between spontaneous speech and read speech appear to lie not in the TYPE of reductions but in the RANGE OF DEGREE of reduction. Read speech seems to run on a sort of cruise control (with non-dramatic readers), so that hyper forms are rare. The need for and sensitivity to information structure in spontaneous speech communication leads to a much more variable rhythmic structure, greater variation in accenting and deaccenting, and the result is not only frequent cases of extreme reduction, but also cases of hyperspeech forms, as examples (33) and (34) illustrate.

(33) Bulgarian ((a) read; (b) spontaneous)
Toj belija dom se namira vav tri santimetra otgore i okolo edin i polovina vav l'avo na lista.
('The white house is three centimetres above and about one and a half to the left on the page.')
(a) [tɔɪ bɛ̯ljəˈdɤ̯m sənə̯ˈmĭᵊrəː . . . fˈtrɪsɐ̯ntʲɨ̯mɛ̯ʈrŭdˈ̚gɔ̯rɛ̯ᵊ . . . ⁱʊkʊədĭnĭpŭːˡʊɪnɐː . . . fˈlʲavŭ . . . nəˈlĭstɐ̯]
(b) [tɔ̯ĕ bɛ̯ljɜˈdɔ̯mːɨ̯ . . . sɐ̯nɨ̯ˈmi̯ᵊrᵊɜ̯ʔə . . . vəf . . . əm . . . ˈtrĭsənt̯ɐ̯mɛtrŏdˈ̚gɔ̯rə̯ . . . ʔɨ . . . ʔɔkŭədĭnʲ əpŭːʊʲɪnɐː . . . ʔɜː . . . vʌ̯f ˈlʲavŭ . . . nəˈlĭstɐ̯]

(34) Russian ((a) read; (b) spontaneous)
Znachit iz starta ja vyxozhu nalevo i plavno kak by podnimajus' naverx.
('Well, from the start, I go to the left and gradually, I sort of go up.')
(a) [ᶻnazˠə̯də̯sˈtɑ̯rtejɑ̯ŭʏĕzˠʊnɑ̯ˈlʲeŭ ĭʊˈplʸɐnʸə̯kɐgɯpᵠdnʲɞmɑ̯ɐ̯snaʸˈʊɪᵊrˣ]
(b) nɐᵈzˠə̯d̥ə̯zˈtɑ̯ːrd̥ĭjɑ̯ɛʊ . . . ʊŭxɐzʊnɑ̯ˈlʲeəw̯ɜᵊ . . . ĭˈplaᵘnʊkɑ̯gᵇˈpⁱdʲmɑ̯ᵠsnăˈvĭərˣ . . .]

The spontaneous utterances in both languages manifest the tempo and phrasing differences, with hesitations and pauses, and consequent pre-pausal lengthening. In both languages, the spontaneous version is considerably longer in total than the read one (Bulgarian: 11.58 sec. vs. 6.17 sec.; Russian: 5.22 sec. vs. 3.95 sec.), although some individual phrases and parts of phrases are in fact shorter. In terms of segmental variation, the Bulgarian spontaneous utterance illustrates a case of non-reduction of a preposition (*vav*: [vəf]) before a hesitation, where the read version has the standard

fully reduced form [f]. On the other hand, in the Russian utterance, the underlying /z/ in the initial /zn/ cluster is completely elided in the spontaneous version, whereas in the read version a partially reduced cluster remains: the canonically fully voiced /z/ is realised as a very weak devoiced [z̥]. In the modal particle *kak by* (sort of, somehow) both versions show a merging of /k/ and /b/ properties: in the read version, the voicing of /b/ and the place of articulation of /k/ remain as a homogeneous phonetic [g] segment preceding the vowel of underlying *by*, whereas the spontaneous version reveals a very brief voiced bilabial closure after a [g] segment before a clearly voiceless /p/ closure, the vowel of *by* having been elided.

In conclusion, the examples of reduction across the languages examined here can be seen as evidence to support the assumptions stated at the outset. Variation in the time and effort invested in different parts of an utterance is a language-independent phenomenon, and the propensity for articulatory reduction must therefore also be language-independent. Despite considerable differences of sound inventory and syllable structure across the languages from a European standpoint, the six languages must be seen as quite similar in terms of the variety in the languages of the world. It is, therefore, perhaps unsurprising that the similarities in reduction were clearer than possible differences. This study should, therefore, be taken as a stimulus to examine more extensive data from the languages represented here and to extend the study of continuous speech, both read and spontaneous in their various forms, to a wider range of languages.

References

Abercrombie, D. (1967). *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.

AIPUK – Arbeitsberichte 30 (1996). *Sound Patterns in Spontaneous Speech*, Kohler, K. J., Rehor, C. & Simpson, A. (eds.). Institut für Phonetik und digitale Sprachverarbeitung, Universität Kiel.

AIPUK – Arbeitsberichte 31 (1996). *Sound Patterns of Connected Speech. Description, Models and Explanation*, Simpson, A. & Pätzold, M. (eds.). Institut für Phonetik und digitale Sprachverarbeitung, Universität Kiel.

AIPUK – Arbeitsberichte 32 (1997). *The Kiel Corpus of Read/Spontaneous Speech: Acoustic Database, Processing Tools and Analysis Results*, Simpson, A. P., Kohler, K., J. & Rettstadt, T. (eds.). Institut für Phonetik und digitale Sprachverarbeitung, Universität Kiel.

Anderson, A. H., Bader, M., Bard, E. G., Boyle, E., Doherty, G., Garrod, S., Isard, S., Kowtko, J., McAllister, J., Miller, J., Sottillo, C., Thompson, H. & Weinert, R. (1991). The HCRC Map Task Corpus. *Language and Speech* **34(4)**, 351–366.

Bolinger, D. (1965). *Pitch Accent and Sentence Rhythm, Forms of English: Accent, Morpheme, Order*. Cambridge, MA: Harvard University Press.

Comrie, B. & Corbett, G. G. (eds.), *The Slavonic Languages*. London: Routledge.

Dasher, R. & Bolinger, D. (1982). On pre-accentual lengthening. *Journal of the International Phonetic Association* **12**, 58–69.

Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* **11**, 51–62.

Dauer, R. M. (1987). Phonetic and phonological components of language rhythm. *XI International Congress of Phonetic Sciences*, vol. 5, 447–450. Tallinn.

Farnetani, E. & Busà, M. G. (1999). Quantifying the range of vowel reduction in Italian. *Proceedings of XIV International Congress of Phonetic Sciences*, 491–494. San Francisco.

Gimson, A. C. (1989). *An Introduction to the Pronunciation of English* (4th edn.), revised by A. Ramsaran. London: Edward Arnold.

Grabe, E. & Low, E. L. (to appear). Acoustic correlates of rhythm class. *Papers in Laboratory Phonology* **6**.

Hardcastle, W. J. & Barry, W. J. (1989). Articulatory and perceptual factors in /l/ vocalisations in English. *Journal of the International Phonetics Association* **15(2)**, 3–17.

Hardcastle, W. J. & Marchal, A. (eds.) (1990). *Speech Production and Speech Modelling*. Dordrecht: Kluwer.
Hoequist, C. E. (1983a). Durational correlates of linguistic rhythm categories. *Phonetica* **40**, 19–31.
Hoequist, C. E. (1983b). Syllable duration in stress-, syllable- and mora-timed languages. *Phonetica* **40**, 203–237.
Hoequist, C. E. (1983c). The perceptual center and rhythm categories. *Language and Speech* **26**, 367–376.
Kohler, K. J. (1979). Kommunikative Aspekte satzphonetischer Prozesse im Deutschen. In Vater, H. (ed.), *Phonologische Probleme des Deutschen*, 13–39. Tübingen: Günther Narr Verlag.
Kohler, K. J. (1990). Segmental reduction in connected speech in German: phonological facts and phonetic explanations. In Hardcastle & Marchal (eds.), 69–92.
Kohler, K. J. (1995). *Einführung in die Phonetik des Deutschen* (2nd edn.). Berlin: Erich Schmidt Verlag.
Lindblom, B. (1990). Explaining phonetic variation: a sketch of the H&H theory. In Hardcastle & Marchal (eds.), 403–439.
Lloyd-James, A. (1940). *Speech signals in telephony*. London.
Nespor, M. (1990). On the rhythm parameter in phonology. In Roca I. M. (ed.), *Logical Issues in Language Acquisition*, 157–175. Dordrecht: Foris.
Pike, K. L. (1945). *The Intonation of American English*. Ann Arbor, MI: University of Michigan Press.
Pointon, G. E. (1980). Is Spanish really syllable-timed? *Journal of Phonetics* **8**, 293–304.
Ramus, F. & Mehler, J. (1999). Language identification with suprasegmental cues: a study based on speech resynthesis. *Journal of the Acoustical Society of America* **105**, 512–521.
Ramus, F., Nespor, M. & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition* **73**, 265–292.
Scatton, E. A. (1975). *Bulgarian Phonology*. Cambridge, MA: Slavica Publishers Inc.
Scatton, E. A. (1993) Bulgarian. In Comrie & Corbett (eds.), 188–248.
Ternes, E. & Vladimirova-Buhtz, T. (1999). Bulgarian. In *Handbook of the International Phonetic Association*, 55–57. Cambridge: Cambridge University Press.
Tilkov, D. & Bojadzhiev, T. (1977). *Bylgarska Fonetika*. Sofia: Nauka i izkustvo.
Timberlake, A. (1993). Russian. In Comrie & Corbett (eds.), 827–886.
Uldall, E. (1972). Relative durations of syllables in two-syllable rhythmic feet in R.P. in connected speech. *Work in Progress* **5**. Department of Linguistics, Edinburgh University, 110–111.
Wenk, B. & Wioland, F. (1982). Is French really syllable-timed? *Journal of Phonetics* **10**, 193–216.