Rhythmic modification in Child Directed Speech

Elinor Payne
Phonetics Lab and St Hilda's College, University of Oxford

Brechtje Post
RCEAL and Jesus College, University of Cambridge

Lluïsa Astruc
The Open University and Dept of Spanish and Portuguese, University of Cambridge

Pilar Prieto
Departament de Traducció i Ciències del Llenguatge, ICREA-Universitat Pompeu Fabra

Maria del Mar Vanrell
Departament de Filologia Catalana, Universitat Autònoma de Barcelona

ABSTRACT: Interval-based rhythm metrics were applied to the speech of English, Catalan and Spanish mothers addressing their children (aged 2, 4 and 6 years) and compared with their speech when addressing other adults. Results show that when mothers speak to their children, their speech is more vocalic (higher %V) and more even-timed (lower variability in vocalic and consonantal interval duration). Little differentiation is made when addressing children of different ages for the window of development investigated here. These results indicate a greater 'syllable-timed' quality to child-directed-speech, irrespective of the presumed rhythmic category of the language in question. Characteristic cross-linguistic rhythmic differences observed in adult-directed-speech are, in the main, preserved, although more weakly, suggesting that while modifications in child-directed speech largely occur within the bounds of a presumed rhythmic category, such categories may, in fact, overlap.

1. Introduction

This paper examines the extent to which mothers modify the rhythmic properties of their speech when addressing their children, in what have traditionally been described as 'stress-timed' and 'syllable-timed' languages. To measure rhythmic properties we apply a set of well known 'rhythm metrics' which extract and quantify certain (predominantly durational) properties from the acoustic signal (cf Ramus *et al.* 1999; Dellwo 2006; Low *et al.* 2000; Grabe-Low 2002; White-Mattys 2007)[1]. While these metrics have been widely applied to capturing perceived cross-linguistic differences in rhythm in (largely non-spontaneous) reading styles of adult speech, little attention has been paid to determining whether such differences are evident in different speech contexts known to trigger different speaking styles. One such different speaking style is that used when addressing small children and infants, hereafter referred to as child-directed-speech (CDS). We examine the rhythmic properties of CDS for English, Spanish and Catalan as a function of child age, and compare them with adult-directed speech (ADS) for the same languages. We consider the degree to which temporal/rhythmic modifications observed in CDS may be interpreted as an exaggeration of rhythmic properties in canonical adult speech or an accommodation towards the speech of the child in question.

1.1. Background on speech rhythm measurement

Though the acoustic basis of perceived cross-linguistic differences is elusive, the percept of a distinction between two broad language categories (i.e. between the *Morse Code* rhythm of e.g. Germanic languages and the *machine gun* rhythm of e.g. Romance languages) is strong, and empirically supported. For example, Ramus *et al.* (2003), show adult ability to distinguish *between* rhythmic categories, but not *within* them. The validity and efficacy of rhythm metrics in capturing perceived differences have been widely discussed and evaluated (see in particular an

1. Following White-Mattys (2007), other rhythm metrics, based on variability in the duration of phonological constituents (e.g. Deterding 2001), were not applied.

overview and comparison of all metrics by White-Mattys 2007), and only a brief outline of their development will be given here.

Empirical studies (cf. Pointon 1980; Lea 1974; Dauer, 1983; Roach 1982) found no evidence to support the early thesis (Pike, 1945; Abercrombie, 1967) that cross-linguistic differences result from top-down timing strategies selecting different units (foot versus syllable[2]) for isochrony. Following the empirical discreditation of the isochrony account, an alternative explanation (see Bertinetto 1981; Dauer 1983; 1987; Roach 1982; Dasher-Bolinger, 1982) proposed that rhythmic distinction emerges from distinct sets of phonological and phonetic properties found across languages, most notably the complexity of syllable structure and the presence versus absence of vowel reduction. In brief, it was observed that in so-called 'stress-timed' languages, like English, there is a greater range of syllable structures (permitting complex codas and onsets), heavier syllables are more likely to attract stress, and unstressed vowels tend to be reduced (both durationally and quantitatively, see Delattre 1966). In contrast, in so-called 'syllable-timed' languages like Spanish, open syllables are far more common, and vowel reduction is much less evident. This hypothesis of rhythmic difference (which we call the 'phonology-derived' hypothesis) in effect captures differences in the degree to which languages use duration to execute prosodic prominence. On this interpretation, the execution of prosodic prominence is not merely dependent on transparent timing strategies (hence the failure to find the basis of rhythm in isochrony) but interacts in a complex manner with the segmental string. To the extent that the proposed rhythm metrics 'work' (i.e. discriminate between perceived differences), it is because they have access to the acoustic effects of this interaction, without having to 'know' the component parts of this interaction (i.e. the language-specific structural and phonetic properties).

These claims find strong support in psycholinguistic studies (Nazzi *et al.* 1998; see also Ramus *et al.* 2003; Nazzi-Ramus 2003; Nazzi *et al.* 2000) showing that infants attend to rhythmic dif-

2. Whence the terms 'stress-timed' and 'syllable-timed'.

ferences from birth. This evidence indicates that cross-linguistic rhythmic differences reflect something that can be perceived before lexical or phonological analysis is available to the infant, and therefore can be captured objectively from the acoustic stream. From this, Ramus *et al.* (1999) argue that "[the] infant primarily perceives speech as a succession of vowels of variable durations and intensities, alternating with periods of unanalysed noise (i.e. consonants)". This insight led Ramus *et al.* (1999) to develop three measures of utterance rhythm which can be extracted purely on acoustic grounds, on the basis of vocalic and consonantal intervals: i) the standard deviation of vocalic intervals (ΔV); ii) the standard deviation of intervocalic (i.e. consonantal) intervals (ΔC); and the proportion of utterance duration which comprises vocalic intervals (%V). Application of these metrics (Ramus *et al.* 1999; Ramus *et al.* 2003) to languages of perceived different rhythmic categories revealed a combination of ΔV and either ΔC or %V to be the most useful in distinguishing categories.

Evidence that ΔC and ΔV are inversely proportional to between-language differences in speech rate (e.g. Barry *et al.* 2003; Dellwo-Wagner 2003; Lee-Todd 2004) led Dellwo (2006) to propose the rate-normalised metric VarcoC, and Ferragne-Pellegrino (2004) and White-Mattys (2007) to propose the rate-normalised metric VarcoV[3]. VarcoC and VarcoV are variation coefficients calculated as the standard deviation of the interval duration in question (ΔC or ΔV) divided by the mean interval duration (mean C or mean V), and multiplied by 100.

A different approach to measuring the basis of rhythmic difference, but one nevertheless derived from acoustic intervals, is that proposed by Low *et al.* (2000), namely the application of the pairwise variability index (PVI). Rather than taking global 'vocalic-ness' and variability in that 'vocalic-ness', the PVI attempts to capture *sequential* differences in vocalic interval duration, and specifically between stressed and unstressed syllables. The

3. VarcoC also varied according to speakers' intended speech rate, though not in a consistent way across the languages in that study (French, German and English).

motivation for looking at the sequential nature of the contrast is that prosodic prominence depends on syntagmatic contrast: what counts is a comparison with what has gone immediately before and with what lies immediately ahead. The PVI is calculated as the mean of the differences between successive intervals, and is normalised (nPVI) for variability of speech rate by dividing by the sum of intervals. Grabe and Low (2002) also propose an intervocalic PVI (PVI-C), but warn against normalising this. The rationale for this is that since the size and variability of intervocalic intervals largely reflect a language's phonotactics and since these are claimed to be an underlying source of that language's rhythmic properties, normalising eliminates rhythmic difference.

With the proliferation of different metrics, the question arises as to how well they perform comparatively. White and Mattys (2007) carried out a direct comparison of the usefulness of rhythm metrics in capturing perceived differences between 'stress-timed' English and Dutch on the one-hand and 'syllable-timed' French and Spanish on the other (for adult speech). They found that %V[4] and the rate-normalised vowel metrics (VarcoV and nPVI-V) to be the most effective (with a slight favouring of VarcoV over nPVI-V, also for within-category discrimination). They failed to find a significant correlation between %V and speech rate (see also Dellwo-Wagner 2003), and it would appear that this metric is fairly robust across different rates. For interval variability metrics, White and Mattys (2007) found that while rate normalisation increased the effectiveness of *vocalic* measures (VarcoV was found to have greater discriminatory power than ΔV), this was not mirrored in the case of consonantal metrics. No evidence was found for systematic patterning between rhythmic classes when using VarcoC. This is compatible with Grabe and Low's advocacy of leaving their intervocalic PVI measure raw (rPVI-C). Thus, normalising speech variability over an utterance is deemed to be helpful for capturing vocalic interval variability, but unhelp-

4. Or %C, since these amount to the same thing.

ful for capturing consonantal interval variability[5]. These findings have informed the choice of metrics for building rhythmic 'profiles' for CDS in the present study. We use rate-normalised metrics for vocalic interval variability (VarcoV and nPVI-V) and non-normalised metrics for consonant interval variability (ΔC and rPVI-C), in addition to the measure of 'vocalic-ness', %V.

1.2. Characteristics of Child Directed Speech

Studies have shown that CDS differs from ADS notably in prosodic patterns and speaking rate (see e.g. Albin-Echols 1996; Fernald 2000; Fernald-Mazzie 1991; Fernald-Simon 1984; Garnica 1977; Grieser-Kuhl 1988), though also in segmental properties (see e.g. Watt *et al.* 2003; Englund-Behne 2006). Common features that have been identified include higher and greater range of pitch, especially in stressed syllables; longer duration of individual words, more prominent final lengthening, slower speech rate, higher amplitude, shorter utterances and longer pauses, and more reliable positioning of pauses at phrase boundaries (Fernald-Simon 1984; Snow-Ferguson 1977; Stern *et al.* 1982, 1983).

Evidence from cross-linguistic studies suggests that higher pitch and wider pitch range are almost universal properties of CDS, British and American English (Fernald *et al.* 1989; Shute-Wheldall 1989; Stern *et al.* 1983), German (e.g., Fernald-Simon 1984; Fernald *et al.* 1989), Italian and French, (Fernald *et al.* 1989), Mandarin Chinese (Grieser-Kuhl 1988; Papousek-Hwang 1991) and Thai (Kitamura *et al.* 2002). However, the extent of exaggeration varies and exceptions exist: Grieser and Kuhl (1988) report less exaggerated pitch characteristics in CDS in Mandarin Chinese (a tonal language) than in American English, and Fernald *et al.* (1989) found no evidence for higher pitch and wider range

5. Elsewhere, however, there is evidence that rate normalised consonant metrics *do* discriminate. For example, Dellwo (2006) reports a higher VarcoC for English and German than for French (though no statistical tests were reported).

in Japanese CDS. Kitamura *et al.* (2002) compare mean F0, pitch range and utterance slope-F0 in CDS in Thai (a tonal language) and Australian English, and find intonational features to be more exaggerated in CDS in both languages, but especially in Australian English.

Evidence also suggests that properties of CDS evolve as a function of the child's development (Fernald 2000; Shockey-Bond 1980). In addition to differences in topic of conversation (Snow 1977), and shifts in the balance between affect and information-salient speech (Sherrod *et al.* 1978), there is evidence for age-related acoustic modifications. For example, pitch height and range in adults' speech have been shown to decrease as the age of the child increases (Garnica 1977; Warren *et al.* 1984; Stern *et al.* 1983: combined, these studies cover the 4mth to 5 year age range).

From a functional point of view, it has been argued that the richness of prosodic cues in CDS may actively function to attract and maintain the attention of the child (Fernald-Simon 1984; Stern *et al.* 1982) and communicate affective information (Fernald 1989; 1992; Werker-McLeod 1989). It has also been argued that CDS facilitates language acquisition by, for example, highlighting linguistic structure and aiding word identification (Fernald-Mazzie 1991; Cooper-Aslin 1989; Fernald 1992; Gleitman *et al.* 1988; Hirsh-Pasek *et al.* 1987; Kemler Nelson *et al.* 1989). In support of this thesis, Kemler Nelson *et al.* (1989) showed that 7- and 10-month old infants were sensitive to prosodic cues to clause boundaries in CDS but not in Adult-Directed Speech (ADS). More generally, research has shown that young children and infants prefer listening to CDS than to ADS (Cooper-Aslin 1990; Fernald 1985; Pegg *et al.* 1992; Werker-McLeod 1989) Infants' preference for CDS also appears to be universal (Werker *et al.* 1994; Fernald 1993). These preferences appear to be mainly based on modifications in pitch (Fernald-Kuhl 1987), rather than in other parameters such as duration. Furthermore, there is evidence (Kitamura-Burnham 1998) that what the children are actually responsive to is relative prominence. Thus a preference for exaggeratedly high pitch is not a preference for a high pitch *per se*, but for the greater contrast with surrounding material. A contrasting low pitch can also provoke a response.

2. Objectives of present study

The present study investigates whether mothers modify the language-specific rhythmic characteristics of their speech when addressing their children. The question is potentially of interest to studies of the development of rhythm and prosody in child's speech, and to phonetic and phonological development more generally. In addition, it may contribute to our understanding of speech in interaction and the parameters along and within which speech patterns can be manipulated. The study also dovetails with questions concerning the basis of perceived differences in rhythm cross-linguistically. As has been detailed above, according to the phonology-driven account, perceived differences in rhythm *between* languages supposedly result from different ways in which these languages break up sequences of vocalic articulations, and are thus closely correlated to phonotactics and other phonological properties such as syllable structure and vowel reduction. Since it would be reasonable to expect adults to maintain syllable structure when changing speech style to address children (i.e. we would not expect them to start simplifying consonant clusters or deleting codas), if we adhere to a strong version of the phonology-driven account, strictly speaking we should not expect to find modification of rhythm in CDS[6].

The desirability of taking a broader approach to understanding the bases of perceived rhythmic difference between languages has been highlighted by several authors. Beckman (1992: 457) stresses the need to "examine the actual lengths of component syllables as a function of their position within the hierarchy of stresses and phrases". White and Mattys (2007: 518) also draw attention to the possible influences of "language-specific prosodic timing processes, such as accentual lengthening, word-initial lengthening and phrase-final lengthening" on interval-based rhythm metrics, and point to the need to examine more fully the degree to which differences in prosodic timing processes correlate with rhythm class

6. Although a less strong interpretation might allow some modification through exaggerated or diminished vowel reduction.

distinctions. In their investigation of rhythmic difference between northern and southern varieties of Italian, White *et al.* (2009) find systematic differences in prosodic timing processes. Prieto, Vanrell, Astruc, Payne and Post (*under revision*) show that differences in the rhythm metrics between English on the one hand and Spanish and Catalan on the other clearly emerge even when syllable structure is controlled for. Furthermore, they show that the rhythmic distinctions between these languages correlate with differences in the way these languages instantiate durational marking of prosodic heads and prosodic boundary lengthening. Thus, while phonotactic properties undoubtedly do affect rhythm metrics, evidence suggests they are not the only, or indeed necessarily the most important, factor in determining rhythmic category. How a language instantiates prosodic timing processes may be as important, possibly more so. Given the known extensive manipulation of prosodic characteristics in CDS, this potentially leaves considerable scope for manipulation of rhythmic category in CDS.

Following on from this position, the first hypothesis that we test is that CDS does indeed differ from ADS in terms of the rhythm metrics. Once we allow for the possibility of modification of rhythmic category within a particular language, the question arises as to what *kind* of modification takes place. The second part of the hypothesis therefore concerns the way in which rhythm could be modified in CDS. Modifications could potentially take various forms. On the basis of previous findings for CDS, which report a strong tendency for adults to exaggerate the prosodic properties of ADS, one possibility is that the rhythmic profiles of ADS are exaggerated in CDS. Since these profiles are language-specific and are said to cluster into more or less well-defined rhythmic categories (e.g. 'stress-' versus 'syllable-timed'), an exaggeration would entail different kinds of modifications for different languages. Dominey and Dodane (2004: 128) claim that "the essential acoustic property of CDS is the exaggeration or modulation of characteristics that are *already present* in ADS"[7]. These

7. Echoed elsewhere in the CDS literature, e.g. Albin and Echols (1996).

characteristics are different for different categories of rhythm. Thus, applied to rhythm, in so-called 'stress-timed' languages, which are characterised by a low proportion of vocalic stretches overall (low %V) but a high variability in both vocalic and consonantal stretches (high VarcoV, nPVI-V, ΔC and rPVI-C), CDS would exhibit *even lower* %V and *even higher* VarcoV, nPVI-V, ΔC and rPVI-C. Conversely, in so-called 'syllable-timed' languages, which are characterised by a high %V but low variability in vocalic and consonantal stretches, CDS would exhibit *even higher* %V and *even lower* VarcoV, nPVI-V, ΔC and rPVI-C. A further consequence of such category-sensitive exaggeration would be that cross-category differences would be greater for CDS than for ADS.

Given the precedence for exaggerating prosodic properties that exist in ADS, our expectation is that CDS will show exaggerated ADS rhythmic characteristics too. However, an alternative possibility is that, rather than exaggerating the characteristics of ADS, adults accommodate towards the temporal/rhythmic characteristics of child speech. Studies on rhythm in child speech have suggested that it starts out uniformly more 'even-timed', i.e. showing characteristics associated with 'syllable-timed' languages and that more 'uneven' timing characteristics are acquired later on, for languages like English (see Allen-Hawkins 1978, 1980; Grabe *et al.* (1999). Grabe *et alii* report that the speech of 4 year-old French and English speakers is more 'syllable-timed' than that of their mothers, and that the difference is particularly marked for English. However, more recent research (see Payne, Post, Astruc, Prieto and Vanrell *under revision*[8]) paints a more complex picture, with the rhythmic properties of child speech failing to fall into any well-defined rhythmic class. In both 'syllable-timed' Spanish and Catalan, and 'stress-timed' English, child speech (for ages 2, 4 and 6 years) was shown to be more 'vocalic' (higher %V) than adult

8. Payne *et alii* (*under revision*) used the same set of subjects as the present study. The child speech described here was taken from the same dialogues as the CDS reported in the present study.

speech in the same language and to have a tendency towards lower variability in vocalic interval duration: these are both indices of greater ‘syllable-timing’. However, consonant interval variability was shown to be *higher* in child speech, which is generally considered to be a property of more ‘stress-timed’ rhythm. A possible explanation put forward by Payne *et alii* for this divergent behaviour of consonants and vowels is that the mastery of interval variability in developmental speech is fundamentally different for vowel and consonant production. It is suggested that variability in vocalic interval duration is largely associated with prosodic structure, the finer points of which are acquired later on than the age range under investigation, while variability in consonant interval duration is largely correlated with immature articulatory co-ordination abilities, still characteristic of speech for the age range under investigation (Allen-Hawkins 1980). Thus, on this interpretation, the task for the child is to introduce vocalic variability but to bring consonantal variability under control. A straightforward accommodation to the child’s speech would therefore entail higher %V and more even timing for vocalic intervals (lower VarcoV and nPVI-V) coupled with less even timing for consonant intervals (higher ΔC and rPVI-C).

Since it has been shown that CDS can be sensitive to the age of the child being addressed, our second hypothesis is that the rhythm metrics will show a difference in scores between speech addressed to 2, 4 and 6 year-olds.

Our third and final hypothesis is that despite any rhythmic modifications made in CDS, cross-linguistic differences observed for ADS will remain detectable also in CDS. This is partly motivated by the reasoning that the underlying phonotactics and syllable structure also remain the same for ADS and CDS within any given language, and that these do contribute to rhythm metric scores.

The languages chosen for this study differ in certain phonological and prosodic properties, and are said to differ according to rhythmic class (for ADS). English and Spanish are the most straightforward to categorise. English (Southern Standard

British variety) is canonically classed as 'stress-timed' in the rhythm literature. It displays a wide variety of syllable structure types, quantitative vowel reduction[9] (Delattre 1966), stress attraction to heavier syllables, and substantial final lengthening (Wightman *et al.* 1992). Spanish (European variety) is canonically classed as 'syllable-timed'. It is dominated by CV syllable structure, and displays weak vowel reduction (Delattre 1966), a weak correlation between stress and syllable weight, and weaker final lengthening.

Catalan holds a more ambiguous position in the schema of rhythmic classification. Though its predominant syllable structure is CV (typical of 'syllable-timed' languages), it has a greater incidence of more complex syllables than Spanish, weak vowel reduction, a moderate correlation between stress and syllable weight, and more final lengthening than in Spanish (Ortega-Llebaria and Prieto 2007). The apparent mix of 'syllable-timed' and 'stress-timed' properties has led some researchers, following a strong interpretation of the 'phonological' view of rhythm, to claim that Catalan is rhythmically 'intermediate' (Nespor 1990; Ramus *et al.* 1999). However, there is contradictory evidence from recent empirical studies. Prieto *et alii* (*under revision*) show that when syllable structure properties are controlled for, no differences arise between Catalan and Spanish in the rhythm metrics. Further evidence for the 'syllable-timed' status of Catalan is given by Gavaldà-Ferré (2007), who shows that different degrees of vowel reduction found in different dialects of Catalan make no difference to rhythm scores. The inclusion of Catalan in the present study was to ascertain whether these moderate differences from Spanish induces different patterns in children acquiring rhythm.

9. Vowel reduction in English, in the form of centralisation to schwa, is of course also qualitative. Such differences – not detectable by duration-based rhythm metrics – may also contribute to the percept of rhythm, but are outside the scope of the present study.

2.1. Summary of hypotheses

We hypothesised that:

H1 CDS would show rhythmic divergence from ADS, and more specifically, would exaggerate the characteristics of ADS rhythm (and therefore diverge from ADS in an unsystematic way cross-linguistically);

H2 CDS scores would show rhythmic divergence as a function of child age;

H3 CDS scores would show some degree of cross-linguistic divergence, as observed in ADS.

3. Method

3.1. Participants

For each language, we recorded twelve mother-child dialogues (to elicit CDS), and made separate recordings of the mothers interacting with an adult interviewer (to elicit ADS). The children fell into three age groups (i.e. there were four children in each age group for each language): 2-, 4- and 6-year-olds. Some mothers were recorded in two different dialogues, with children of different ages. In total, 36 children and 26 adults were recorded[10].

3.2. Materials and elicitation

The data consisted of short question-and-answer dialogues, elicited through the medium of a structured game, based on short, animated clips, shown on Powerpoint slides on a laptop screen. The animations showed simple, everyday scenes, which could easily be described in words that were highly familiar to the children.

10. This study was part of a wider study which also investigated rhythm in child speech.

For example, one scene showed a little girl blowing soap bubbles, another showed a little boy playing with building blocks.

The mother was instructed to ask her child to describe what was happening in each clip, then praise the child for getting it right, and repeat what the child had said. The mothers were also recorded doing the same task, but interacting with an adult (the interviewer). A typical dialogue went thus:

CDS context	
Mother:	*What's happening here? What's the little girl doing?*
Child:	*"(She's) blowing bubbles!"*[11]
Mother:	*That's right!* ***<u>She's blowing bubbles</u>****!*
ADS context	
Mother:	*What's happening here? What's the little girl doing?*
Interviewer:	***<u>She's blowing bubbles</u>****!*
Mother:	*That's right! She's blowing bubbles!*

The utterances used for the present study were the target utterances spoken by the mother (in the example above the instances of *She's blowing bubbles* in bold and underlined). It is to be noted that this form of elicitation for CDS depends on a repetition of what the child has said. In theory, this may make it more likely for the adults to adopt mimicry strategies and accommodate to characteristics present in the child speech. However, this could also happen even when an utterance is not repeated but simply a new utterance in the context of a larger dialogue. It is difficult, and arguably contrived to elicit CDS that is 'immune', so to speak, to the online effects of hearing child speech, and it is arguably precisely such interaction which induces a change in speech style.

Tab. 1 shows the number of utterances produced by mothers in dialogue with each age group, and in dialogue with other adults, for each language. Spanish CDS utterances directed at 6-year olds are missing: this is because, in this particular recording session, the mothers were mistakenly not instructed to repeat what their child had said.

11 Or an approximation.

Table 1.

	CDS 2-year olds	CDS 4-year olds	CDS 6-year olds	CDS total	ADS
English	69	77	68	214	226
Catalan	69	64	68	201	154
Spanish	65	33	-	98	138

Recordings were made respectively in the participants' homes in Cambridge, Madrid and Barcelona, using a Marantz PMD660 recorder and Shure PG81 microphones for the Spanish and Catalan recordings, and a Tascam HD-P2 recorder with AKG C3000B microphones for the English recordings.

3.3. Analysis

Vocalic intervals were segmented and labelled (start-points and end-points) from the waveform and spectrogram by a research assistant using *Praat* (Boersma-Weenink 2007). The RA is a native speaker of English and studied Spanish language as part of her degree. The Catalan and Spanish labellings were revised by the three author native speakers. Vocalic and consonantal segmentation was performed according to standard criteria, e.g. F2-onset for vowels (Peterson-Lehiste 1960). For instance, in fricative-vowel sequences, the onset of the vowel was taken to be the beginning of the second formant. The onset of a fricative was marked at the start of high frequency energy. Prevocalic glides were considered to be part of the consonantal intervals; postvocalic glides part of the vocalic intervals. The following metrics were calculated (after Grabe-Low 2002 and White-Mattys 2007):

– Measure of 'vocalic-ness': %V
– Measures of variability in vocalic interval duration (rate-normalised): VarcoV and nPVI-V
– Measures of variability in consonantal interval duration (not rate-normalised): ΔC and rPVI-C

Rate-normalised metrics were chosen for vocalic intervals and non-normalised metrics were chosen for consonant intervals since, as discussed in Section 1, these have been shown to have greater discriminatory power for adult speech in general.

In addition, speech rate was calculated by dividing the number of vocalic intervals (roughly equivalent to the number of syllables) by the total time in seconds of each utterance. A repeated measures analysis, with Speaker Style and Language as factors, showed there was no main effect of Speaker Style. However, there was a significant interaction of Speaker Style*Language ($F(2, 402)=13.961$, $p<0.001$). As can be seen in Fig. 1[12], showing mean speech rate for ADS and CDS in all three languages, speech rate is marginally slower in CDS than in ADS for Spanish and Catalan, but is faster for English.

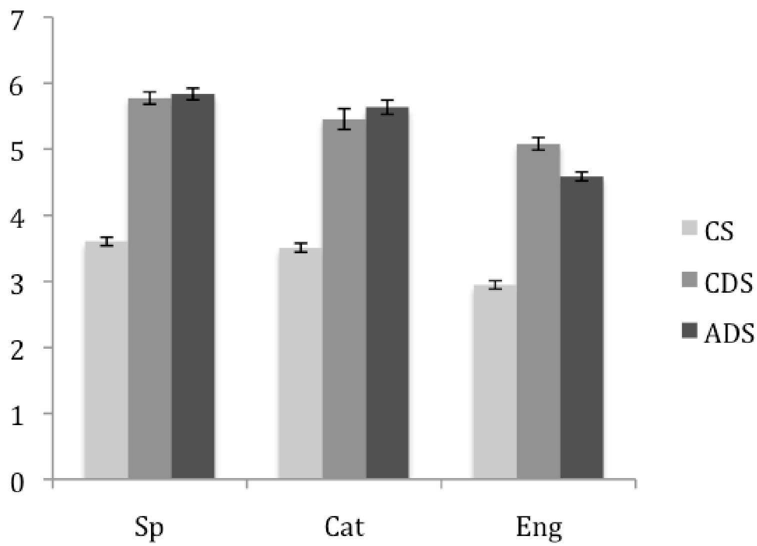

Figure 1. Mean speech rate in vocalic intervals/second. Bars represent one standard error around the mean

12. Fig. 1 also shows mean speech rate for child speech. The present study formed part of a larger study also looking at child speech.

4. Results and Discussion

Since the semi-spontaneous elicitation technique used in this study differs from most previous studies on adult speech rhythm (which in the main elicited data through subjects reading prepared sentences), as a first step, we report briefly on the robustness of the cross-linguistic differences for ADS, to verify that our elicitation technique replicates previous findings. Our findings for ADS are reported in more detail in a related study investigating the rhythmic characteristics of child speech (see Payne *et alii under revision*). The graphs in Fig. 2 and Fig. 3 are reproduced from that paper, and plot, respectively %V against Varco V and nPVI-V against rPVI-C. All of the metrics discriminated very robustly between English on the one hand and Spanish/Catalan on the other. English showed greater variability in both consonant and vocalic intervals (higher VarcoV, nPVI-V, ΔC and rPVI-C) and was less 'vocalic' (lower %V). These results reproduce a

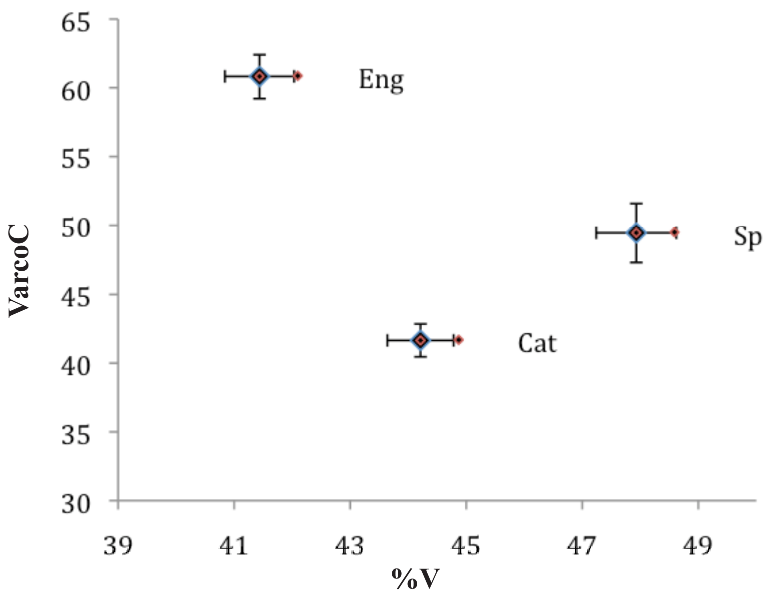

Figure 2. Distribution of English, Spanish and Catalan over the %V, VarcoC plane. Bars represent one standard error around the mean

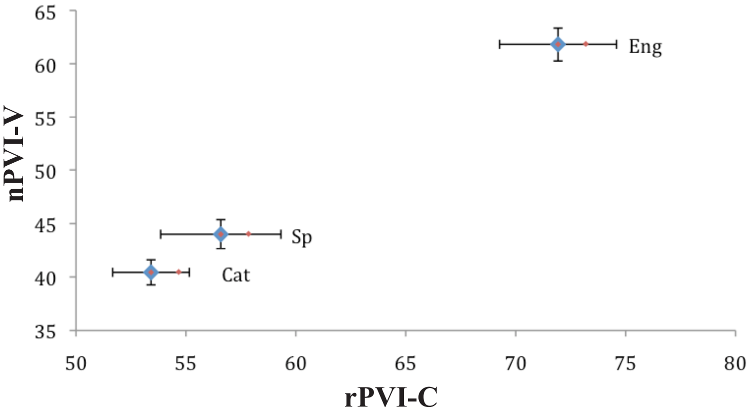

Figure 3. Distribution of English, Spanish and Catalan over the nPVI-V, rPVI-C plane. Bars represent one standard error around the mean

very similar pattern to the results reported by White and Mattys (2007) for English and Spanish. The distinction between Catalan and Spanish was much weaker, and not systematic in terms of rhythmic class. Catalan was found to be less vocalic (lower %V) – indicating greater 'stress-timing', but less variable in vocalic interval duration (lower VarcoV) – indicating greater 'syllable-timing'.

The results on CDS are reported in the following order: i) rhythmic characteristics of CDS compared with ADS; ii) CDS sensitivity to child age; iii) cross-linguistic differences in CDS.

4.1. Hypothesis 1: Rhythmic divergence of CDS

Our first hypothesis was that CDS rhythm scores would differ from ADS scores, and specifically that they would exaggerate the rhythmic properties of ADS in the language concerned. A multivariate Anova was run on the CDS and ADS scores, with Speaker Style and Language as factors. For each section, we report results on a) vocalic-ness (%V); b) variability in vocalic interval duration (VarcoV and nPVI-V); and c) variability in consonant interval duration (ΔC and rPVI-C). Graphs plotting %V against VarcoV and rPVI-C against nPVI-V, for each language (following White and Mattys, 2007) are given at Fig. 4 through 9.

4.1.1. Vocalic-ness (%V)

There was found to be a main effect of Speaker Style on %V ($F(1, 1031) = 8.31$, $p < 0.05$), with %V being significantly higher in CDS than ADS. However, there was a significant interaction of Speaker Style*Language ($F(1, 1031) = 17.06$, $p < 0.001$), and Bonferroni post hoc comparisons showed that the difference between CDS and ADS was only significant for English ($p < 0.001$) and Catalan ($p < 0.05$).

4.1.2. Vocalic interval variability (nPVI-V and VarcoV)

Comparing ADS and CDS, for vocalic interval variability, there was found to be a main effect of Speaker Style on the global measure of variability, VarcoV ($F(1, 1006) = 535.9$, $p < 0.001$) with variability significantly *lower* in CDS than in ADS. There was no interaction of Speaker Style*Language, and post hoc comparisons showed a significant difference for all three languages ($p < 0.001$ for all languages). However, for the sequential measure of variability, nPVI-V, there was no main effect of Speaker Style, and no interaction of Speaker Style*Language.

4.1.3. Consonant interval variability

There was found to be a main effect of Speaker Style on the global measure of variability ΔC ($F(1, 1029) = 23.97$, $p < 0.001$), with variability significantly *lower* in CDS than in ADS. There was no interaction of Speaker Style*Language, and post hoc comparisons revealed that differences between CDS and ADS were significant for all languages ($p < 0.001$ for each language). For the sequential measure (rPVI-C), there was no main effect of Speaker Style, and no interaction of Speaker Style*Language.

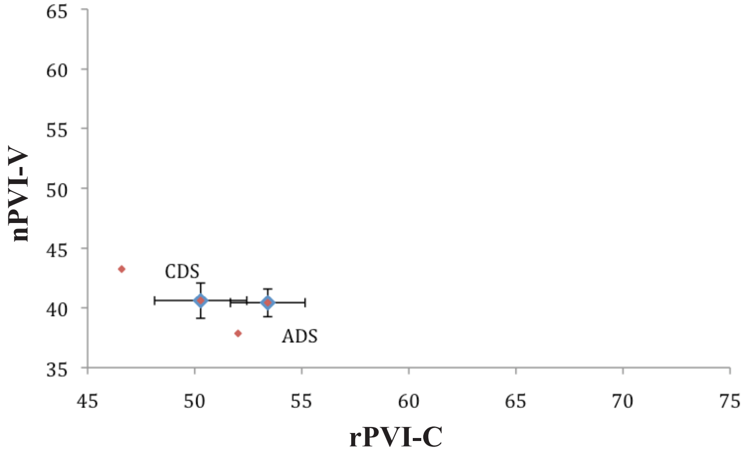

Figure 4. Distribution of CDS and ADS in Catalan over the rPVI-C, nPVI-V plane. Bars represent one standard error around the mean

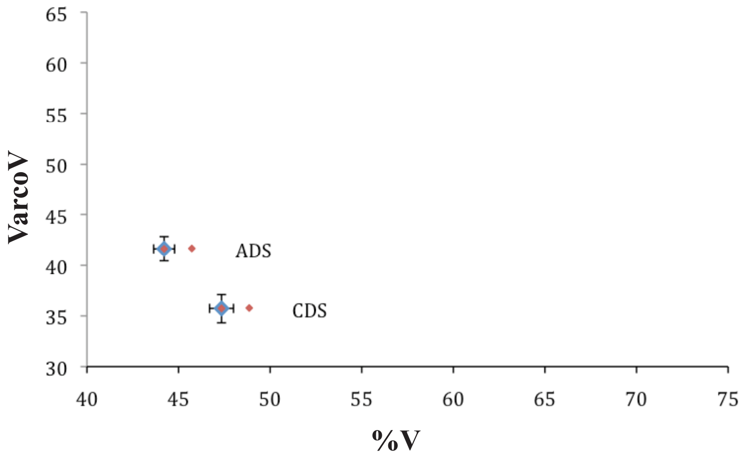

Figure 5. Distribution of CDS and ADS in Catalan over the %V, VarcoV plane. Bars represent one standard error around the mean

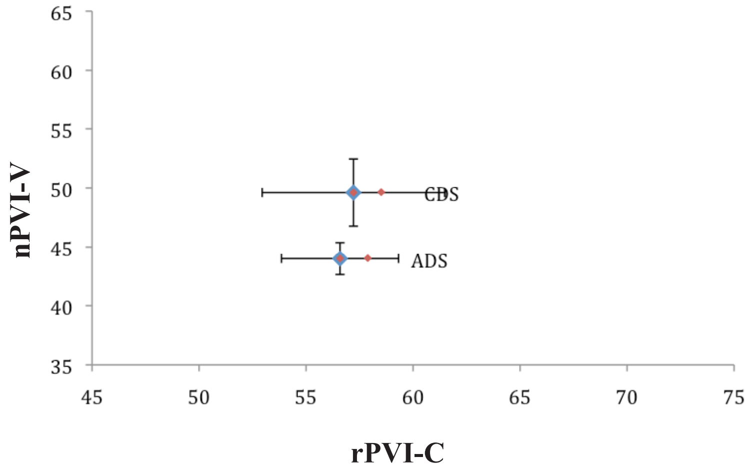

Figure 6. Distribution of CDS and ADS in Spanish over the rPVI-C, nPVI-V plane. Bars represent one standard error around the mean

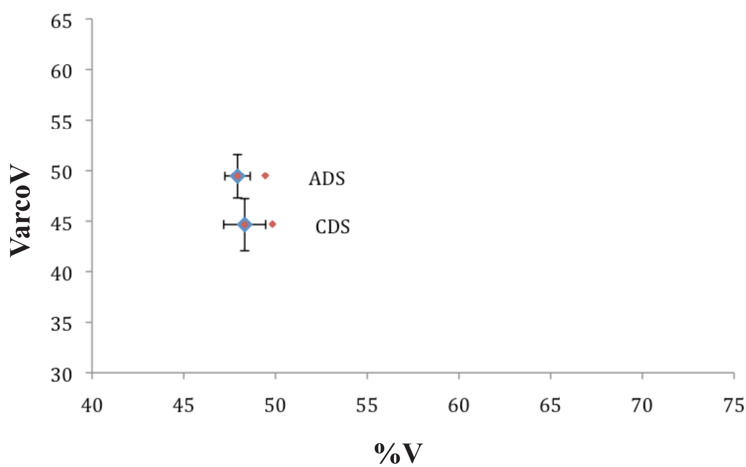

Figure 7. Distribution of CDS and ADS in Spanish over the %V, VarcoV plane. Bars represent one standard error around the mean

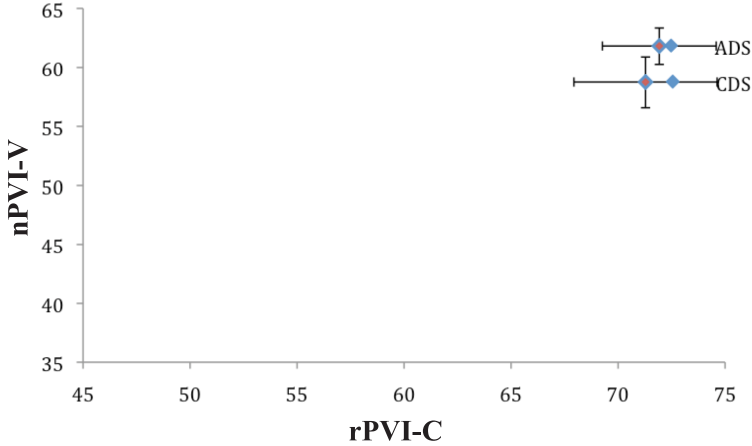

Figure 8. Distribution of CDS and ADS in English over the rPVI-C, nPVI-V plane. Bars represent one standard error around the mean

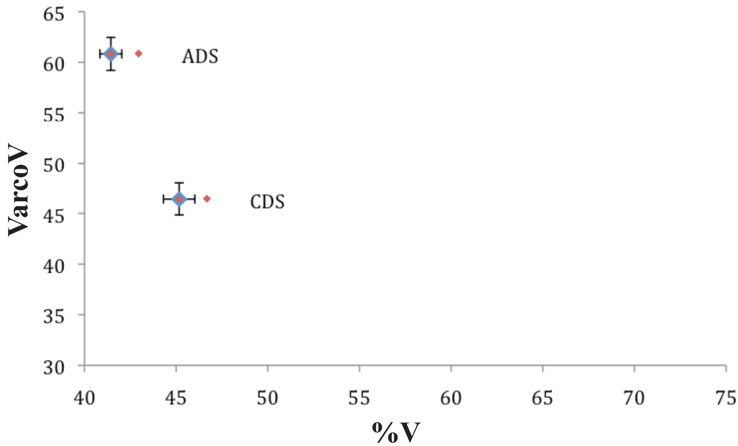

Figure 9. Distribution of CDS and ADS in English over the %V, VarcoV plane. Bars represent one standard error around the mean

4.1.4. Discussion of Hypothesis 1

The results show that the rhythmic characteristics of CDS do differ from those of ADS, though not uniformly across all measures and across the languages examined. Cross-linguistically, the most robust distinctions were found using the global measures of variability, (rate-normalised VarcoV for vowels and non-normalised ΔC for consonants), with scores being significantly lower in CDS than in ADS. %V was also found to be significantly higher in CDS than in ADS for Catalan and English. Thus mothers show less *global* variability in both consonant and vowel interval duration, and had more vocalic speech when speaking to their children than when speaking to the adult interviewer. These results point to CDS being more 'even-timed' and showing more characteristics of 'syllable-timing' than ADS, particularly in Catalan and English, but also in Spanish. Hence, rhythmic modification in CDS appears to induce the bases of a greater 'syllable-timed' effect, irrespective of the rhythmic category of the language concerned. While this could be interpreted as an exaggeration of ADS characteristics for Catalan and Spanish, the same cannot be said of English.

The sequential measures of variability, i.e. the rate-normalised metric for vowels (nPVI-V) and the non-normalised metric for consonants (rPVI-C) did not discriminate between CDS and ADS for any of the three languages. Why global and sequential measures should differ in this respect is open to speculation, but could be attributed to the different types of variability that the two kinds of measure are designed to detect. Sequential measures (PVI) are sensitive to consecutively alternating patterns, and will be highest when the rhythm is an approximation of 'stressed-unstressed-stressed-unstressed' etc., with frequently occurring reduced/weak syllables. According to the 'exaggeration' hypothesis, CDS in a stress-timed language such as English would be expected to exaggerate the duration of stressed syllables and reduce the duration of unstressed syllables even further, thereby increasing the sequential variability. The results clearly go against this hypothesis. Global measures (Δ and Varco) on the other hand, do not distinguish between consecutive and non-consecutive variability, and are therefore more sensitive to variability across different parts of the utterance. Since the global measures showed less variability in CDS, we might speculate that when adults address young children they adopt a more consistent tempo throughout the utterance. Again, the lower variability seen in English CDS goes against the 'exaggeration' hypothesis.

4.2. Hypothesis 2: CDS sensitivity to child age

Our second hypothesis was that the rhythm metrics would discriminate between CDS directed towards children of different ages. A series of univariate ANOVAs were run on CDS scores with Child Age and Language as factors.

The results showed that there was no main effect of Child Age on %V, VarcoV, ΔC or rPVI-C. The only metric for which there was a main effect of Child Age was nPVI-V ($F(2, 488) = 3.142$, $p < 0.05$). Bonferroni post hoc comparisons showed that the only significant difference was between age 2 and 4 ($p < 0.001$), with nPVI-V scores being lower when mothers were addressing 2 year-olds. There was no significant interaction of Child Age*Language for nPVI-V.

4.2.1 Discussion of Hypothesis 2

In the main, the results do not support the hypothesis that rhythmic modification in CDS is sensitive to child age, with the possible exception of mothers reducing vocalic variability even more when addressing the youngest children in the study. Since CDS has been shown to differ from ADS, it remains for further investigation to establish towards which age of child adults begin adopting ADS-like rhythmic characteristics in their speech.

4.3. Hypothesis 3: Cross-linguistic differences in CDS

Our third hypothesis was that CDS would show the same cross-linguistic divergence in rhythmic scores observed in ADS.

The mean scores for nPVI-V and rPVI-C cross-linguistically are given in Fig. 10, while those for %V and Varco are given in Fig. 11. A series of univariate Anovas were run on the CDS scores, with Language as a factor. The results support this hypothesis to some degree: as with ADS, there was a main effect of Language for all metrics (%V: $F(2, 510) = 3.28$, $p < 0.05$; VarcoV: $F(2, 487) = 9.40$, $p < 0.001$; nPVI-V : $F(2, 488) = 20.16$, $p < 0.001$; ΔC: $F(2, 508) = 7.48$, $p < 0.001$; rPVI-C: $F(2, 508) = 12.45$, $p < 0.001$). Bonferroni post hoc comparisons showed that, as for ADS, English has higher variability than Catalan for VarcoV ($p < 0.001$), nPVI-V ($p < 0.001$), ΔC ($p < 0.001$) and rPVI-C ($p < 0.001$). For %V, English has lower mean scores than Catalan, but since $p = 0.059$, this difference is just above the level of significance. Unlike ADS, English is only distinct from Spanish for rPVI-C ($p < 0.05$) and nPVI-V ($p < 0.05$) (in both cases English has higher variability). Catalan CDS is distinct from Spanish CDS only in the vocalic variability metrics VarcoV ($p < 0.05$) and nPVI-V ($p < 0.05$), and in both cases has lower variability (as with ADS).

Although the variability metrics clearly discriminate between English and Catalan for CDS, the discrimination on the basis of 'vocalic-ness' appears to be somewhat undermined by speech style modifications. This is true also of the comparison between English and Spanish CDS. %V is reported to be a very robust discriminator of rhythmic class cross-linguistically, suggesting that it reflects a fundamental structural difference underlying the percepts of 'stress-' and 'syllable-timing'. It is therefore perhaps somewhat surprising that adult speech can be modified in such a way as to obscure this distinction. The distinction between Spanish and English, also very clear in ADS, is even more obscured in CDS than the distinction between English and Catalan, with only sequential measures of variability being effective discriminators. On the other hand, the distinction between Catalan and Spanish is consistently weak in ADS and CDS.

The first part of this study has shown that CDS in all three languages exhibits modifications characteristic of greater 'syllable-timing'. Having established this, the fact that the difference between 'stress-timed' English and 'syllable-timed' Catalan and (especially) Spanish is diminished in CDS, but not the difference between Spanish and Catalan, would suggest that the greater degree of modification is undergone by English. In other words, all three languages appear to move further along a hypothetical rhythmic continuum in the direction of greater ‘syllable-timing’, but English moves further along, relative to its usual location for ADS.

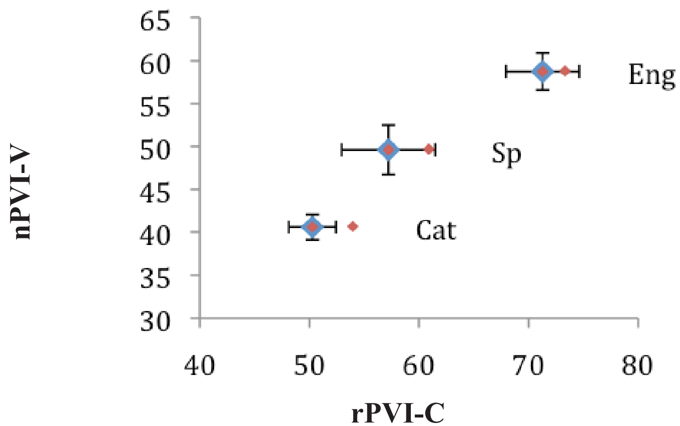

Figure 10. Distribution of Catalan, Spanish and English CDS over the rPVI-C, nPVI-V plane. Bars represent one standard error around the mean

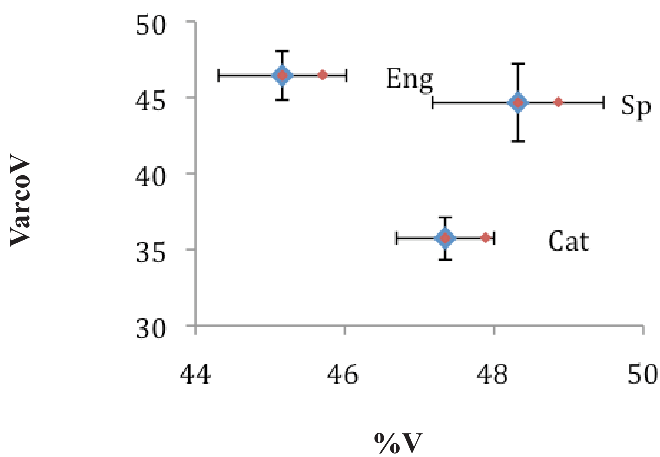

Figure 11. Distribution of Catalan, Spanish and English CDS over the %V, VarcoV plane. Bars represent one standard error around the mean

5. General Discussion

A clear finding of this study is that while rhythmic class distinctions detectable in ADS continue to be observable in CDS, they are less robust. English and Catalan remain well discriminated by variability scores. However, the evidence for English having a lower %V is weak, at best. The distinction between English and Spanish is considerably weaker than in ADS : only the *sequential* variability scores successfully discriminate between them. Thus, there appears to be a degree of blurring in the categorisation of what, traditionally, have been described as canonically 'stress-timed' and 'syllable-timed' languages for this style of speech. At the same time, a very weak distinction between Catalan and Spanish observed in ADS remains. These modifications in cross-category differences appear to be brought about by a general tendency towards greater 'syllable-timing' in CDS. Cross-linguistically, mothers increase the vocalic proportion and decrease the interval variability (both vocalic and consonantal) of their speech when they address infants. The change to consonant metrics is particularly interesting since it has been claimed by some (cf Grabe and Low 2002) that consonant interval variability is an artefact of a language's phonotactic properties. The fact that adults vary this for different speaking styles in the same language would suggest that this parameter of timing *can* be controlled for speech style.

What exactly is underpinning the modifications in CDS, whether they are indicative of exaggeration or accommodation, is hard to determine. For Spanish and Catalan, a shift further towards the more 'syllable-timed' end of a hypothetical rhythm continuum could be interpreted as an exaggeration of properties observed in ADS for these languages. However, this certainly cannot be the case for English. The mothers investigated in the study were possibly accommodating, to some degree, to the speech of their children, which has been shown (for the same children) to be more "syllable-timed" along certain parameters than the ADS target for all languages. However, as has been reported in Payne *et alii* (*under revision*), child speech is not more 'syllable-timed' along *all* parameters: with respect to consonant variability child speech has more uneven timing, i.e. appears to be more 'stress-timed'. In this respect, the lesser consonant variability in CDS alone appears to be an exaggeration of a characteristic of ADS.

The rhythmic profile of CDS is, therefore, neither a wholesale accommodation to the rhythmic profile of child speech, nor a wholesale exaggeration of ADS rhythm. Nevertheless, CDS strikingly displays the characteristic of more 'syllable-timed' (when compared with ADS) across rhythmic categories of languages. Positioning CDS between child speech on the one hand and ADS on the other, cross-linguistically, we can describe the commonalities as follows: vocalic-ness and vocalic variability appear to accommodate towards child speech, while consonant variability appears to be more similar to ADS (indeed exaggerates this characteristic).

How might we plausibly seek to explain this divergence? A divergence in the behaviour of vowels and consonants in child speech has already been highlighted (Allen-Hawkins 1978, 1980; Payne *et alii under revision*). In the study by Payne *et alii*, the transition towards the adult speech target, for both 'syllable-' and 'stress-timed' languages, vocalic variability is interpreted as something *systematic but not yet acquired*, while consonantal variability is interpreted at least in part[13] as something *relatively random and to be eliminated/reduced*. By reducing both consonant and vocalic variability (an overall strategy of greater 'syllable-timing', adults are effectively adapting vocalic production for accommodating effect and adapting consonant production for possible 'instructive effect'. On this interpretation, the rhythmic profile of CDS would be characterised by incomplete accommodation. A possible alternative (but keeping accommodation in consonants and instruction in vowels) would be to introduce more *uneven* timing all round. However, it would be arguably a pointless and more difficult task for

13. This part being variability due to immature phonetic abilities. More systematic differences in consonant variability are due to phonotactic differences.

adults to introduce spurious and random consonant variability, than to strip their speech of some of the finer prosodic nuances, nuances which are arguably being indexed anyway through exaggerated pitch modulation.

6. Conclusions and next steps

The interval-based metrics applied in this study yield a picture of weakened cross-linguistic rhythmic distinctions (weaker than those previously established for ADS). The fact that this weakening is even brought about, and in natural, adult speech (albeit in specific stylistic context), raises the question as to whether we can neatly categorise languages according to rhythm. At the very least, it would appear that some degree of elasticity within categories needs to be allowed, allowing scope for fluctuations in the way speech in a particular language can be delivered, and even the possibility that categories may overlap.

At the same time, the metrics reveal cross-linguistic similarities in the way CDS differs from ADS: CDS is more even-timed than ADS, irrespective of the rhythmic category to which the language in question is said to belong. The fact that, from the production perspective, temporal characteristics deemed at least to *contribute* to the percept of rhythm vary systematically between two styles of adult speech within a given language, and that supposedly robust distinctions between languages can be much reduced, shows that these characteristics do not emerge from phonotactic and syllable structure properties alone. This adds strength to the claim (cf. Beckman 1992; White-Mattys 2007; White *et al.* 2009; Prieto *et alii under revision*) that the rhythmic percept is not entirely, perhaps not even principally, dependent on such properties, and that a full model of speech rhythm is one that incorporates prosodic structure and the phonetic instantiation of this. Such an expanded model would also accommodate individual, speaker-specific parameters of rhythm, in addition to the style-specific parameters such as the one investigated in the present study.

It should be borne in mind that the differences reported here are in acoustic indices and not perceptual judgments. It has indeed been shown that these indices correlate strongly with a cross-linguistic perceptual distinction in ADS, whence the claim that they actually form, or help form, the underlying *basis* of this distinction. However, we know very little about how these indices map onto the rhythmic percept and we do not know – and this is critical – the *degree* to which they must vary in order to trigger the percept of rhythmic difference[14]. The indexical differences reported here could prove perceptually weak or incoherent. Thus a next step will be to test the perceptual validity of the differences found. In particular, it will be of interest to determine whether English and Spanish CDS are perceived as rhythmically different, since the indexical distinction between these is much weaker in CDS than in ADS. Another priority will be to determine whether CDS is perceived as rhythmically different from ADS for the same language, especially by infants, and, following on from that, whether infants show any preference for CDS rhythm.

Finally, that the speech context investigated here is shown to exhibit a degree of similarity cross-linguistically is of particular interest, since it implies that cross-linguistic patterns in the manipulation of speech can and do exist. The degree to which adults consciously manipulate their speech patterns, and for what purpose, is open to conjecture, and we offer some thoughts for further speculation here. In the case of more even timing in CDS, this is more likely to be an artefact of the adult trying to speak more clearly, rather than a conscious effort to modify rhythm, and specifically to accommodate her speech patterns to those of the child. One way of establishing this would be to compare the

14. Other potential dimensions to the perception of rhythm, such as intensity, vowel quality and F0, and how these may be integrated with durational differences, lie beyond the scope of this paper, but remain to be explored more fully. In one of the few studies of this kind, Lee and Todd (2004) report evidence that 'stress-timed' English and Dutch show greater variability in vocalic intensity than 'syllable-timed' Italian and French, for ADS.

CDS rhythmic profile with that of consciously careful and clear ADS. This would inform us as to whether it is *generally* the case that speakers reduce vocalic variability when they adopt a more careful, emphatic style, or whether this is something peculiar to the style of speech adopted when addressing infants. If the former, this would mean that the rhythmic modification observed in CDS is not actually an accommodation to the rhythm of child speech *per se*, so much as an accommodation to the *needs* of the child (which happen to coincide, in part). If the latter, this would presumably mean that the accommodation is actually to certain formal properties of the child's speech. Though the result is the same, the underlying aim would be subtly different. A closer analysis of the interaction between pairs of mother and child could also throw light on what is driving the modification of speech in CDS, and will form a next step for the investigation of the speech corpus used in this study.

Acknowledgements

This research has been funded by a Batista i Roca research project entitled *The acquisition of rhythm in Catalan, Spanish and English* (Ref.: 2007 PBR 29), awarded by the *Generalitat de Catalunya*, with Pilar Prieto as PI. We are grateful to Núria Argemí, Anna Barberà, Ana Estrella and Melanie Bell for recording the data in different languages, and especially to Naomi Hilton for performing the segmentation of the data. Many thanks to Greg Kochanski for helpful comments on an earlier draft of this paper, and to Bill Wells and other colleagues who provided helpful discussion of related presentations at BAAP 2008 and EASR 2008.

References

Abercrombie, D. (1967) *Elements of general phonetics. Edinburgh:* Edinburgh University Press.

Albin, D. D. / Echols, C. H. (1996) "Characteristics of stressed and word-final syllables in infant-directed speech: Implications for word-level segmentation", *Infant Behavior and Development* 19: 401-418.

Allen, G. D., / S. Hawkins (1978) "The Development of Phonological Rhythm", in A. Bell / J. B. Hooper (eds.) *Segments and Syllables*, Amsterdam: North-Holland, 172-185.

Allen, G. D. / S. Hawkins (1980) "Phonological Rhythm: Definition and Development", in G. H. Yeni-Komshian / G. H. Kavanagh, J. F / C. A Ferguson (eds.) *Child Phonology*, vol. 1: *Production*. New York: Academic Press, 227-256.

Barry, W. J. / Andreeva, B. / Russo, M. / Dimitrova, S. / Kostadinova, T. (2003) "Do rhythm measures tell us anything about language type?", in M.-J. Solé / D. Recasens / J. Romero (eds.) *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona 3-9 August 2003, Barcelona: Causal Productions Pty Ltd, 2693-2696.

Beckman, M. E. (1992) "Evidence for speech rhythms across languages", in Y. Tohkura / E. Vatikiotis-Bateson / Y. Sagisaka (eds.) *Speech Perception, Production and Linguistic Structure,* Oxford: IOS Press, 457-463.

Bertinetto, P.M. (1981) *Strutture Prosodiche dell'Italiano,* Firenze: Accademia della Crusca.

Boersma, P. / Weenink, D. (2007) *Praat: doing phonetics by computer (Version 5.1.12) [Computer program]*. Retrieved August 4, 2009, from http://www.praat.org/

Cooper, R. P. / Aslin, R. N. (1989) "The language environment of the young infant: Implications for early perceptual development", *Canadian Journal of Psychology* 43: 247-265.

Cooper, R. P. / Aslin, R. N. (1990) "Preference for infant-directed speech in the first month after birth", *Child Development* 61: 1584-1595.

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", in E. Magno Caldognetto / L. Tonelli / K. Vagges / P. Cosi *Proceedings of the 11th International Congress of Phonetic Sciences,* Tallinn 1-7 August 1987, Tallinn, Estonia: Academy of Sciences of the Estonian S.S.R., vol. 5: 447-450.

Delattre, P. (1966) "A comparison of syllable length conditioning among languages", *International Review of Applied Linguistics in Language Teaching* 4: 183-198.

Dellwo, V. (2006) "Rhythm and speech rate: A variation coefficient for deltaC", in P. Karnowski / I. Szigeti (eds.), *Language and language processing,* Frankfurt am Main: Peter Lang.

Dellwo, V. / Wagner, P. (2003) "Relations between language rhythm and speech rate", in M.-J. Solé / D. Recasens / J. Romero (eds.) *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona 3-9 August 2003, Barcelona: Causal Productions Pty Ltd, 471-474.

Deterding, D. (2001) "The measurement of rhythm: A comparison of Singapore and British English", *Journal of Phonetics* 29: 217-230.

Dominey, P. F. / Dodane, C. (2004) "Indeterminacy in language acquisition : the role of child directed speech and joint attention", *Journal of Neurolinguistics* 17: 121-145.

Englund, K / D. Behne (2006) "Changes in infant directed speech in the first six months", *Infant and Child Development* 15.2: 139-160.

Fernald, A. (1985) "Four-month-old infants prefer to listen to motherese", *Infant Behavior and Development* 8: 181-195.

Fernald, A. / Taschner, T. / Dunn, J. / Papousek, M. /de Boysson-Bardies, B. / Fukui, I. (1989) "A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants", *Journal of Child Language* 16: 477-501.

Fernald, A. (1989) "Intonation and communicative intent in mother's speech to infants: Is the melody the message?", *Child Development* 60.6: 1497-1510.

Fernald, A. (1992) "Meaningful melodies in mothers' speech to infants", in H. Papousek / U. Jurgens / M. Papousek (eds.) *Nonverbal vocal communication: Comparative and developmental approaches*, Cambridge: Cambridge University Press, 262-282.

Fernald, A. (1993) "Approval and disapproval: Infant responsiveness to vocal affect in familiar and unfamiliar languages", *Child Development* 64: 657-667.

Fernald, A. (2000) "Speech to Infants as Hyperspeech: Knowledge-Driven Processes in Early Word Recognition", *Phonetica* 57: 242-254.

Fernald, A. / Kuhl, P. (1987) "Acoustic determinants of infant preference for motherese speech", *Infant Behavior and Development* 10: 279-293.

Fernald, A. / Mazzie, C. (1991) "Prosody and focus in speech to infants and adults", *Developmental Psychology* 27.2: 209-221.

Fernald, A. / Simon, T. (1984) "Expanded intonation contours in mothers' speech to newborns", *Developmental Psychology* 20.1: 104-113.

Ferragne, E. / Pellegrino, F. (2004) "A comparative account of the suprasegmental and rhythmic features of British English dialects", in

Proceedings of MIDL 2004 (*Modelisations pour l'Identification des Langues)*, *Identification des langues et des variétés dialectales par les humains et par les machines*, Paris 29-30 November, Paris: École Nationale Supérieure des Télécommunications, http://www.ddl.ish-lyon.cnrs.fr/fulltext/Ferragne/Ferragne_2004_midl.pdf.

Garnica, O. (1977) "Some prosodic and paralinguistic features of speech to young children", in C.E. Snow / C.A. Ferguson (eds.) *Talking to children: Language input and acquisition*, Cambridge, MA: Cambridge University Press, 63-88.

Gavaldà-Ferré, N. (2007) *Vowel reduction and Catalan speech rhythm*. Unpublished MA Thesis, University College London.

Gleitman, L. R. / Gleitman, H. / Landau, B. / Wanner, E. (1988) "Where learning begins: initial representations for language learning", in F. J. Newmeyer (ed.), *Linguistics: The Cambridge survey*, vol. 3, Cambridge: Cambridge University Press.

Grabe, E. / Low, E. L. (2002) "Durational variability in speech and the rhythm class hypothesis", in C. Gussenhoven / N. Warner (eds.), *Papers in Laboratory Phonology* 7, Berlin: Mouton de Gruyter, 515-546.

Grabe, E. / Watson, I. / Post, B. (1999) "The acquisition of rhythm in English and French", in J.J. Ohala, / Y. Hasegawa / M. Ohala / D. Granville / A.C. Bailey (eds.) *Proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco 1-7 August 1999, San Francisco: University of California and Berkeley, 1201-1204.

Grieser, D. L. / Kuhl, P. K. (1988) "Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese", *Developmental Psychology* 24.1: 14-20.

Hirsh-Pasek, K. / Kemler Nelson, D. G. / Jusczyk, P. W. / Wright Cassidy / K. / Druss, B. / Kennedy, L. (1987) "Clauses are perceptual units for young infants", *Cognition* 26: 269-286.

Kemler Nelson / D.G., Hirsh-Pasek, K. / Jusczyk, P.W. and Cassidy, K.W. (1989) "How the prosodic cues in motherese might assist language learning", *Journal of Child Language* 16: 55-68.

Kitamura, C. / Burnham, D. (1998) "The infant's response to vocal affect in maternal speech", in C. Rovee-Collier (ed.) *Advances in Infancy Research*, vol. 12: 221-236.

Kitamura, C. / Thanavishuth, C. / Burnham, D. / Luksaneeyanawin, S. (2002) "Universality and specificity in infant-directed speech: Pitch modifications as a function of infant age and sex in a tonal and non-tonal language", *Infant Behavior & Development* 24:372-392.

Lea, W. A. (1974) *Univac Report PX10791. Prosodic aids to speech recognition: IV. A general strategy for prosodically-guided speech understanding,* St. Paul, Minnesota: Sperry Univac.

Lee, C. S. / Todd, N. P. M. (2004) "Towards an auditory account of speech rhythm: application of a model of the auditory 'primal sketch' to two multi-language corpora", *Cognition* 93: 225-254.

Low, E.L. / Grabe, E. / Nolan, F. (2000) "Quantitative characterisations of speech rhythm: 'syllable-timing' in Singapore English", *Language and Speech* 43: 377-401.

Nazzi, T. / Bertoncini, J. / Mehler, J. (1998) "Language discrimination by newborns : Towards an understanding of the role of rhythm", *Journal of Experimental Psychology: Human Perception and Performance* 24: 756-766.

Nazzi, T. / Jusczyk, P.W. / Johnson, E.K. (2000) "Language discrimination by English learning 5-month-olds: Effects of rhythm and familiarity", *Journal of Memory and Language* 43: 1-19.

Nazzi, T. / Ramus, F. (2003) "Perception and acquisition of linguistic rhythm by infants", *Speech Communication* 41: 233-243.

Nespor, M. (1990) "On the rhythm parameter in phonology", in I. M. Roca (ed.) *Logical issues in language acquisition*, Dordrecht: Foris, 157-175.

Ortega-Llebaria, M. / Prieto, P. (2007) "Disentangling stress from accent in Spanish: production patterns of the stress contrast in deaccented syllables", in P. Prieto / J. Mascaró / M.-J. Solé (eds.) *Segmental and prosodic issues in Romance Phonology*, Amsterdam/ Philadelphia: John Benjamins, 155-176.

Papousek, M. / Hwang, S. C. (1991) "Tone and intonation in Mandarin babytalk to presyllabic infants: comparison with registers of adult conversation and foreign language instruction", *Applied Linguistics* 12: 481-504.

Payne, E. / Post, B. / Astruc, L. / Vanrell, M. (under revision) "Measuring Child Rhythm", *Language and Speech*.

Pegg, J. E. / Werker, J. F. / McLeod, P. J. (1992) "Preference for infant-directed over adult-directed speech: evidence from 7-week-old infants", *Infant Behavior and Development* 15: 325-345.

Peterson, G. E. / Lehiste, I. (1960) "Duration of syllable nuclei in English", *Journal of the Acoustical Society of America* 32: 693-703.

Pike, K. (1945) *The intonation of American English.* Ann Arbor: University of Michigan Press.

Pointon, G.E. (1980) "Is Spanish really syllable-timed?", *Journal of Phonetics* 8: 293-304.

Prieto, P. / Vanrell, M. / Astruc, L. / Payne, E. / Post, B. (*under revision*) “Prosodic temporal organisation and speech rhythm. Evidence from Catalan, English and Spanish”, *Journal of Phonetics.*
Ramus, F. / Dupoux, E. / Mehler, J. (2003) “The psychological reality of rhythm classes: Perceptual studies”, in M.-J. Solé / D. Recasens / J. Romero (eds.) *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona 3-9 August 2003, Barcelona: Causal Productions Pty Ltd, 337-342.
Ramus, F. / Nespor, M. / Mehler, J. (1999) “Correlates of linguistic rhythm in the speech signal”, *Cognition* 73: 265-292.
Roach, P. (1982) “On the distinction between ‘stress-timed’ and ‘syllable-timed’ languages”, in D. Crystal (ed.) *Linguistic controversies*, London: Edward Arnold.
Sherrod, K. B. / Crawley, S. / Petersen, G. / Bennett, P. (1978) “Maternal language to prelinguistic infants: semantic aspects”, *Infant Behavior and Development* 1: 335-345.
Shockey, L. / Bond, Z. S. (1980) “Phonological processes in speech addressed to children”, *Phonetica* 37.4: 267-274.
Shute, B. / Wheldall, K. (1989) “Pitch alternations in British motherese : Some preliminary acoustic data”, *Journal of Child Language* 16: 503-512.
Snow, C. E. (1977) “The development of conversation between mothers and babies”, *Journal of Child Language* 4: 1-22.
Snow, C. E. / Ferguson, C. A. (1977) *Talking to children : Language input and acquisition*, Cambridge: Cambridge University Press.
Stern, D. / Spieker, S. / Barnett, R. K. / MacKain, K. (1983) “The prosody of maternal speech: infant age and context related changes”, *Journal of Child Language* 10: 1-15.
Stern, D. / Spieker, S. / MacKain, K. (1982) “Intonation contours as signals in maternal speech to prelinguistic infants”, *Developmental Psychology* 18: 727-735.
Watt, D. / G. Docherty / P. Foulkes (2003) “First accent acquisition: a study of phonetic variation in child-directed speech”, in M.-J. Solé / D. Recasens / J. Romero (eds.) *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona 3-9 August 2003, Barcelona: Causal Productions Pty Ltd, 1959-1962.
Warren-Leubecker, A. / Bohannon, J. N. (1984) “Intonation patterns in child-directed speech: Mother-father differences”, *Child Development* 55: 1379-1385.
Werker, J. F. / McLeod, P. J. (1989) “Infant preference for both male and female infant-directed talk: a developmental study of attention-

al affective responsiveness”, *Canadian Journal of Psychology* 43: 230-246.

Werker, J. F. / Pegg, J. E. / McLeod, P. J. (1994) “A cross-language investigation of infant preferences for infant-directed communication”, *Infant Behavior and Development* 17: 323-333.

White L. / S. Mattys (2007) “Calibrating rhythm: First language and second language studies”, *Journal of Phonetics* 35: 501-522.

White, L. / Payne, L. / Mattys, S. L. (2009) “Rhythmic and prosodic contrast in Venetan and Sicilian Italian”, in M. Vigário / S. Frota / M. J. Freitas (eds.) *Phonetics and Phonology*, Amsterdam, Philadelphia: John Benjamins, 137-158.

Wightman, C. W. / Shattuck-Hufnagel, S. / Ostendorf, M. / Price, P. (1992) “Segmental durations in the vicinity of prosodic phrase boundaries”, *Journal of the Acoustical Society of America* 91: 1707-1717.