Cad.Est.Ling., Campinas, (43):93-108, Jul./Dez. 2002

HOW LONG IS A STRESS GROUP?

MICHAEL L. O'DELL & TOMMI NIEMINEN[1]
(University of Tampere, Finland)

RESUMO Num estudo recente, Barbosa sugere que, para uma síntese de fala com naturalidade, o número de sílabas no grupo acentual poderia ser determinado dinamicamente a partir das relações momentâneas dos osciladores acentual e silábico. Consideramos aqui as ramificações desta idéia para entender padrões na fala natural, utilizando e estendendo um modelo de osciladores acoplados hierarquizados previamente desenvolvido.

1. INTRODUCTION

Speech exhibits various rhythms manifested by repetitions of many different kinds. Many of these form hierarchical cycles—faster, lower level cycles repeating within slower, higher level cycles. One example of this, which has received some attention in the literature, is the case of syllables within a stress cycle. In previous work (O'Dell & Nieminen, 1998; 1999; 2001) we have had some success interpreting empirical data in terms of an abstract mathematical model of coupled oscillators. Very typically, for instance, average syllable duration diminishes somewhat while the time elapsing between consecutive stresses (interstress interval, or ISI) grows when the number of syllables per ISI increases, a feature which the model shows is in fact very general for systems of hierarchically coupled oscillations (see also Barbosa, 2000 for a thorough review and additional data). In the present article, inspired by a suggestion in a recent article (Barbosa, 2001), we consider the dynamics of stress group length itself in such a model.

1.1. Coupled oscillator model

The basic idea of the coupled oscillator model is to assume the existence of subrhythms which would exhibit simple oscillatory behavior if observed in isolation. When oscillators are combined into larger systems so that they influence each other, the resulting patterns of rhythm may be much more complex than those of the component oscillators. In some cases, enough is known about the mechanisms underlying a

[1] The authors wish to thank Plinio Barbosa for his comments on an earlier version of this paper.

particular behavior, that detailed models of component oscillators and the ways they influence each other (coupling) may be attempted. In many other cases the mechanisms leading to rhythmic behavior are not understood in detail, or can only be guessed at. Fortunately, however, much of the macroscopic behavior of systems of oscillators is relatively insensitive to the exact details of the oscillators or the couplings involved. A mathematical technique called APD theory (for *averaged phase difference*) has been developed which is abstract enough to derive qualitative conclusions about collections of oscillators in spite of minimal knowledge of the details of the components (cf. Kopell 1988). The essence of this technique is twofold. First, any descriptions of oscillating subsystems are reparameterized in coordinates of phase relative to the system's own limit cycle attractor, or "natural oscillation", reducing the variables involved to phase. If no previous physical description is available we may assume this transformation has already been applied and start with a simple phase description. Operating on its own, such a subsystem will be characterized by

$$\dot{\theta} = \omega \qquad (1)$$

that is, the derivative (or rate of change) of the oscillator's phase (θ) is a constant (ω) expressing the oscillator's "natural" rhythm or eigenfrequency. The next step is to consider the interaction of two (or more) such oscillators, each with its own eigenfrequency. Even with the above simplification, this interaction could in general be a complicated function of the phases of each of the subsystems, but a further simplification is utilized in APD theory. For each subsystem the effects at each phase *difference* are averaged over an entire cycle, giving a simple characterization of the total system in terms of constant eigenfrequencies (ω) along with couplings dependent only on phase differences. In the case of syllables and stress groups, we need coupling functions that depend on *n*, the number of syllables per stress group. Each oscillator will have its own eigenfrequency which we designate θ_1 for the stress group oscillator and θ_2 for the syllable oscillator. We assume the coupling influences may be expressed as a function of a quantity

$$\phi_n = \theta_2 - n\theta_1 \qquad (2)$$

with *n* the number of syllables per stress group. If we further assume that the two coupling influences in opposite directions are identical in form (or close enough that averaging out the differences does not radically change behavior) but opposite in sign, varying only in "relative strength," we arrive at the following system:

$$\begin{aligned} \dot{\theta}_1 &= \omega_1 + H(\phi_n) \\ \dot{\theta}_2 &= \omega_2 - rH(\phi_n) \end{aligned} \qquad (3)$$

where *r* indicates the relative strength (or dominance) of the stress group over the syllable. To find an equilibrium solution, we set the time derivative of ϕ_n to zero:

$$\dot{\phi}_n = (\omega_2 - n\omega_1) - (r+n)H(\phi_n) = 0 \qquad (4)$$

which gives

$$H(\phi_n) = \frac{\omega_2 - n\omega_1}{r+n} = h \qquad (5)$$

for the value of $H(\phi_n)$ at equilibrium. The period of the stress group oscillator (e.g. the time from stress to stress, or ISI) at such an equilibrium (if it exists) can then be calculated as a function of n:

$$T_1(n) = \frac{1}{\omega_1 + h} = \frac{r}{r\omega_1 + \omega_2} + \frac{1}{r\omega_1 + \omega_2} n \qquad (6)$$

The period is thus a linear function of n of the form $I = a + bn$ used by Eriksson (1991). If a and b in Eriksson's formula are estimated empirically by regression analysis, then the relative strength parameter r of equation (3) can be estimated as a/b.

1.2. Barbosa's synthesis model

In his description of a synthesis model implemented for Brazilian Portuguese (BP) Barbosa states

This phrasal stress is considered to be genuinely periodic but the exact location of its pulses is locally modified by higher-level linguistic input constraints (lexical and syntactic-semantic information). (Barbosa, 2001: 969)

and

Following analyses from BP data, the second [ie. stress group] oscillator starts with a fixed period across speech rates (the phrase stress oscillator period). The ratio between this value and the period of the first [ie. syllable] oscillator is then computed and rounded. This new value represents the number of V-to-V units to the next phrasal stress. This number is readjusted in order to coincide with a lexical stressed syllable. (op. cit.)

In terms of our model, we take this to be a suggestion that the parameter n involved in the coupling (equations 2, 4, 5 and 6 above) should be considered a dynamic variable whose value is influenced by system dynamics as well by outside control.

Another intriguing feature of Barbosa's synthesis is the fact "As during speech production mechanism, the generation includes the emergence of silent pauses. – – [I]f at a particular position in the sentence (normally corresponding with phrasal stress beat) the corresponding delivered V-to-V normalized duration is greater than a critical value, the insertion of a silent pause is considered." (op. cit.) In future we hope to investigate the applicability of this suggestion to Finnish. However, in the present article we concentrate on the dynamics of stress group length in our Finnish language corpus.

1.3. Hypotheses

Considering stress group length n from the point of view of dynamic systems theory, we assume that the system comes every now and again (normally once for each stress group) to a bifurcation[2] point when a choice of n is made. If we want to express this

[2] By bifurcation, we mean that a stable situation becomes unstable, forcing the system to make a choice. Prigogine defines bifurcation as "The branching of a solution into multiple solutions as a system

more neutrally we may say that a bifurcation is reached from which many paths lead, each path corresponding to or resulting in a different number of syllables being executed during the next stress cycle. Using Haken's terminology (cf. eg. Haken, 1983), *n* functions as an order parameter. After the bifurcation point we assume the path taken has a potential well deep enough (compared to diffusion) so that changing *n* in the middle of a stress group is a rare event. We might say that all possible values of *n* compete with each other at this bifurcation point. It is obvious from the start that many factors affect the chances for a particular *n* to be realized. Perhaps the most obvious is the lexical and syntactic structure of the language which dictates that some syllables (which we shall call lexically stressable syllables, LSS) are much more likely than others to receive stress and start a new stress group. The question we are concerned with is whether or not the choice of *n* is affected (also) by the coupling of stress and syllable rhythms.

A priori there is good reason to expect a possible influence of coupling on choice of *n* given the coupled oscillator model, since the coupling "force" needed to synchronize stress and syllable oscillators will increase as *n* deviates from $\Omega = \omega_2/\omega_1$, the ratio of eigenfrequencies of the two oscillators. For instance, if $\Omega = 3$, then no coupling at all will be necessary if there are exactly 3 syllables per stress group, but the more the syllable count deviates from this ideal, the more coupling force is needed to maintain synchrony.

2. METHODS

2.1. Corpus

Stress in Finnish is almost always realized on the first syllable of a word and is therefore not lexically distinctive, but rather has a delimitative function. The current corpus, originally analyzed by Nieminen in 1996, consists of a 10-minute Finnish language radio speech broadcast in Finland on Saturday 24th October 1992. It was prepared ahead of time and read aloud, and represents a fairly formal speaking style. Stressed syllables were determined in Nieminen's study auditorily by two trained phoneticians working initially independently, and then conferring in cases of disagreement. (A more detailed account of stress judgments can be found in Nieminen 1996.) The duration of the stress groups varies between 183 and 1903 msec, median 751 msec. The length of stress groups measured in syllables varies from 1 to 12 with a median of 4 syllables. At least 75% are six or less syllables long.

---

parameter is varied." (Prigogine 1997: 201). For more examples of applying dynamic systems theory to language cf. eg. Port & van Gelder 1995.

Table 1: Description of corpus.

	freq	%
stressed	317	21.4
unstressed	1164	78.6
TOTAL	1481	100.0

2.2. Experiment I

It is possible to estimate the probability of a syllable being stressed, given no other information about the syllable. In the present corpus there were 317 stressed syllables out of a total 1481 syllables giving a maximum likelihood estimate of 317/1481 = 0.214045 for the probability in question. However, before we go on to consider the possible effect of length of stress group, we should take into consideration the fact that some syllables are a priori almost impossible to stress. This is why in Barbosa's synthesis model, "This number is readjusted in order to coincide with a lexical stressed syllable" (see above).

For the purposes of the present analysis, we classified all syllables in the corpus into two groups, lexically stressable (LSS) and lexically unstressable (LUS). This classification was done independently of the acoustic analysis according to the following simple rule (cf. Karlsson, 1983): the first syllable of every word was considered stressable including the first syllable of each component of a compound word (i.e. words which are written without an intervening space in Finnish orthography, mainly a noun modifying a following noun, e.g. *kriisipesäke* 'roughly: center of crisis'). Naturally we do not claim that all syllables classified "stressable" in this way are equally likely to receive stress. A more rigorous analysis of lexical, syntactic, semantic, pragmatic and other factors could certainly produce a finer distinction in a priori "stressability." For instance the conjunction *ja* 'and' is not as likely to be stressed (though it is stressed occasionally in our corpus) as the first syllable of a noun such as *sota* 'war'. The relevance of this classification into lexically stressable and unstressable can be checked in the light of our corpus by asking how much more likely (given no additional information) it is for a lexically stressable syllable to receive stress compared to a lexically unstressable syllable. In the present corpus 316 lexically stressable syllables received stress out of a total of 559 lexically stressable syllables giving a maximum likelihood estimate of 0.565 for the probability that a lexically stressable syllable is realized with stress. Compare this to one occurrence of stress on a LUS[3] out of a total of 922 LUS giving a maximum likelihood estimate of 0.001 for the probability that a lexically unstressable syllable is realized with stress. The difference is so clear that statistical tests are hardly necessary. In what follows we will restrict attention to the behavior of LSS. (Of course in a much larger corpus with enough

[3] The single occurrence of stressed LUS in our corpus came on the second to last syllable of the very last word (*'virsikirjas'tamme* "from our hymn book"), possibly emphasizing the end of the speech.

occurrences of stresses on LUS it would be of great interest to examine their distribution as well.)

Table 2: Frequency of stress in LSS and LUS.

	stressed		unstressed		TOTAL
	freq	%	freq	%	freq
LSS	316	56.5	243	43.5	559
LUS	1	0.1	921	99.9	922
TOTAL	317	21.4	1164	78.6	1481

The next question to ask is whether the size of the stress group *n* which would result if a syllable is stressed has any effect on the probability of that syllable being stressed. *n* will of course be the number of syllables which have occurred since the previous realized stress. The natural null hypothesis to test against is that there is a constant probability of stress being realized (on a LSS) regardless of how many syllables have occurred since the last stressed syllable. To test this hypothesis we divide all LSS in the corpus into groups according to *n*, interpreted now as the number of syllables since the previous realized stress occurring in the same breath group, ie. with no intervening pauses. Cases which do not have a previous stress in the same breath group are ambiguous and are left out of the analysis.

Table 3: Frequency and estimated probability of stress, given *n*.

n	total	stressed	*p*
1	17	2	0.118
2	86	24	0.279
3	77	36	0.468
4	66	38	0.576
5	47	34	0.723
6	26	32	0.889
7	15	11	0.733
8	7	4	0.571
9	2	2	1.000
10	3	3	1.000
11	1	1	1.000

This allows us to calculate a separate probability estimate (i.e. the proportion of realized stresses) for each *n* occurring in the corpus. The results of this calculation are shown in Table 3 and in Figure 1. Figure 1 also shows a bar indicating the exact 95% credibility interval for the proportion at each *n* (calculated assuming a binomial distribution for the number of stressed syllables), as well as a dotted line showing the proportion for all cases combined, which is the maximum likelihood estimate for *p* given that *p* is constant across all *n*.

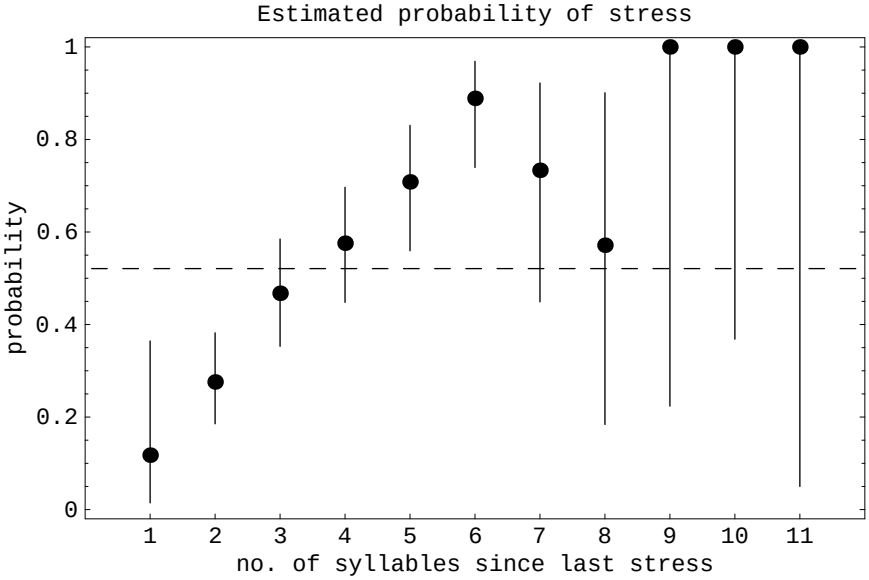

Figure 1: Estimated probability of stress, given *n*.

The relevance of *n* can be tested by considering a Markov chain model with a series of states corresponding to the system being confronted with the choice of stressing or not stressing, i.e. the occurrence of a LSS. To test the relevance of *n*, we consider a system which has a separate LSS state for each *n*, thereby allowing for the possibility that the transition probability is different for different *n*. A complete Markov chain model would include states describing the probabilities of a LSS occurring at various points in the process, but we consider the process only when a LSS has occurred. We utilize a Pearson chi-square statistic for the test (cf. Bhat, 1972:99), lumping together classes $n = 8$ through $n = 11$ to ensure the expected number of cases in each cell is greater than 5. The resulting value of chi-square with 7 degrees of freedom is 31.4804, with a significance of $p < 0.0001$. It would appear we can reasonably reject the hypothesis that choice of stressing a LSS is independent of the number of syllables since the last stress. Indeed, looking at Figure 1, it would seem apparent that there is a clear trend, at least for $n \leq 6$, for stress to be more and more likely the further we get from the last occurrence of stress. (For $n \geq 7$, the estimated probability is extremely unreliable, being based on very few cases, as can be seen by the large credibility intervals in Figure 1.)

2.4 Experiment II

Because the null hypothesis can be rejected, we may try to estimate the form of the probability distribution of *n*, interpreted now as the number of syllables in the upcoming stress group. One way to do this would be to look at the raw empirical frequency of stress groups of various lengths. The problem with this, however, is that we don't know the effect on stress group length which is merely a result of various patterns of LSS

occurring in the data (and presumably in Finnish in general). For instance, it might be the case that the most likely stress group length is four syllables simply because every fourth syllable is most often "available" for stressing (LSS), which could in turn be the result of Finnish lexical structure, in which case the high frequency of four syllables per stress group would not be related to coordination of stress and syllable rhythms at all. In general we might expect there to be some interaction between "availability for stressing" on the one hand and "preference for *n* syllables to a stress group" on the other.

Here we are interested in estimating the probability distribution for choosing between stress group sizes, controlling for the pattern of "available stress positions." To do this we tally the number of cases observed in the corpus for each possible size of stress group *n* for each possible pattern of LSS up to a certain (finite) number of syllables *m*. For definiteness we restrict cases to patterns at the end of their breath group (i.e. ending in a break) so that we can be relatively sure that no LSS farther away than m affected the choice of stress group size. The total number of such patterns will be 2^{m-1}. For instance, let "S" stand for a syllable with realized stress, "U" for a lexically unstressable syllable, and "L" stand for a lexically stressable syllable (whether actually stressed or not). The pattern "SUULUU#" (which occurred fairly often, 14 times in our data) presents a choice between only two stress group lengths: $n = 3$ if stress is realized on the "L", or $n = 6$ if not. The fact that in our data 7 of these cases had a stressed "L" and 7 did not then provides some evidence, ceteris paribus, that choices $n = 3$ and $n = 6$ are similar in regard to probability.

To estimate the probabilities for each *n* and each pattern (a nuisance parameter which we are not directly concerned with) we can fit a log-linear model of independence for a table with structural zeroes (cf. Agresti, 1990), sometimes called a quasi-independence model (a structural zero at some cell in our table corresponds to a lexically unstressable syllable). It was convenient to stop the count at $m = 9$ because thus far there occurred examples of every stress group length in the data, but not for $m = 10$ (that is there were no occurrences of a breath group ending with a stress group 10 syllables long). Thus in what follows, probability estimates should be interpreted as probability of choosing n, given that $n \leq 9$. In addition, instead of tabulating a complete 9×256 table ($m = 9$ syllables by $2^{m-1} = 256$ possible patterns), we leave out all patterns which provide no choice (only one stress group length possible), as well as all patterns which never occurred in our data. While these cases provide some information on the relative popularity of the respective patterns, they provide no information on the probabilities of stress group length. This procedure resulted in a total of 52 patterns left for model fitting.

Since the resulting table was sparse with many (empirical) zeroes, we resorted to Bayesian estimation of parameters using Markov Chain Monte Carlo methods (cf. e.g. Gilks *et al*. 1996) computed with the WinBUGS 1.3 program (Spiegelhalter *et al*. 1996). The specification of the model using BUGS syntax is given in the appendix. For the estimation of parameters we used 100000 iterations after a "burn-in" of 10000 iterations. Convergence was monitored using the Gelman-Rubin statistic (available in BUGS) on two parallel simulation chains.

The resulting estimated distribution is shown in Figure 2. For each *n* the mean value for estimated probability of choosing *n* syllables to a stress group when possible (and given that $n \leq 9$) is shown, along with vertical bars indicating 95% credibility intervals around the means. We note, for instance, that categories $n = 3$ and $n = 6$ are indeed similar in probability (see above discussion). In general it would seem that the various *n* are not equally probable and that there is a preference for stress groups in the mid range around $n = 5$.

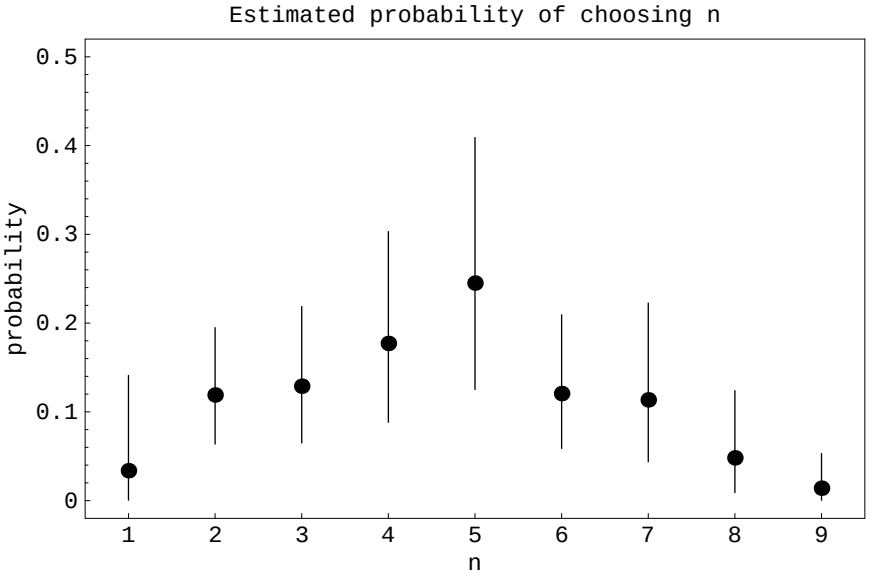

Figure 2: Estimated probability of choosing *n*.

It can be seen in Figure 2 that the 95% CI are fairly wide. This is due of course to the small size of our sample. Because the location of the distributional mode is of considerable importance, we use the BUGS run to estimate the posterior probability that the mode is actually at $n = 5$ and to construct credibility intervals. These probabilities are shown graphically in Figure 3.

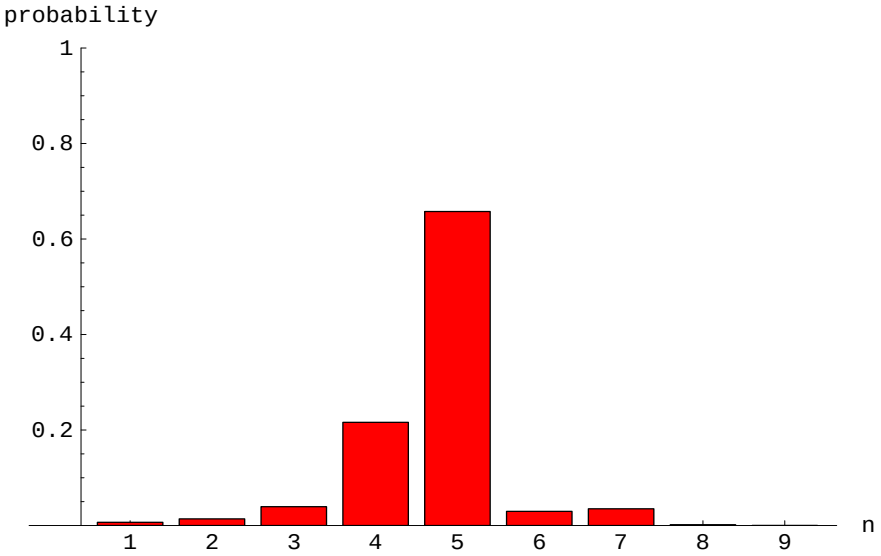

Figure 3: Probability of mode at *n*.

Based on these probabilities CI can be constructed as shown in Table 4. In other words $n = 5$ is most likely to be the mode ($p = 0.666$), and we can be quite certain at least that the mode falls somewhere between $n = 3$ and $n = 7$ ($p > 0.95$). The fact that one category out of nine has such a high posterior probability of being the mode also allows us to reject the hypothesis that choice of stress group length is uniformly distributed, i.e. that any stress group length is equally likely once the effect of stress availability is controlled for.

Table 4: Probability of mode within the credibility interval (CI).

CI	5	4–5	3–5	3–6	3–7
probability	0.666	0.884	0.920	0.942	0.976

3. DISCUSSION

3.1. Influence of coupling on stress group length

It would appear that even with our meager amount of data there is some support for Barbosa's proposal that the number of syllables in a stress group is influenced by the requirements of keeping stress and syllable rhythms synchronized. In Barbosa's synthesis model synchronized stress and syllable oscillators provide a way to choose the most natural placement of stresses given a choice between several lexical stresses. We have shown that similar forces may well be at work in natural speech production as well.

3.2 Relevance for the coupled oscillator model

When there is enough data to give a reliable picture of the mode of the probability distribution for stress group length, that mode may shed some light on the possible values for the eigenfrequencies of the component oscillators. Having "factored out" the lexical effects, it should be the case that the most likely n is close to the ratio of eigenfrequencies $\Omega = \omega_2/\omega_1$. To illustrate this further, we start with the theoretical value of coupling force h needed to keep two hierarchically coupled oscillators synchronized. From equation (5) above we have

$$h(n) = \frac{\omega_2 - n\omega_1}{r + n} \quad (7)$$

Instead of using this value to find the periods of the oscillators as in equation (6), we use it directly as an indication of effort needed for synchronization, considered as a function of n. Allowing n to be dynamically determined instead of fixed we may construct a plausible potential well for n by assuming the derivative of the potential function $\mathrm{d}V(n)/\mathrm{d}n$ is proportional to $h(n)$, giving:

$$V(n) = A(n\omega_1 - (r\omega_1 + \omega_2)\log(r + n)) + C \quad (8)$$

with A and C arbitrary scaling constants which can be set to 1 and 0 respectively without loss of generality.

This formulation is equivalent to saying that the system tends towards values of h which are small in absolute value. Figure 4 shows two examples of such a plausible potential function for two example sets of values for the eigenfrequencies ω_1 and ω_2 and the relative coupling strength parameter r. The first curve (labeled A) has values ω_1 = 2.8 Hz, ω_2 = 4.2 Hz, r = 2.0, while the second curve (labeled B) has values ω_1 = 1.3 Hz, ω_2 = 5.8 Hz, r = 0.9.

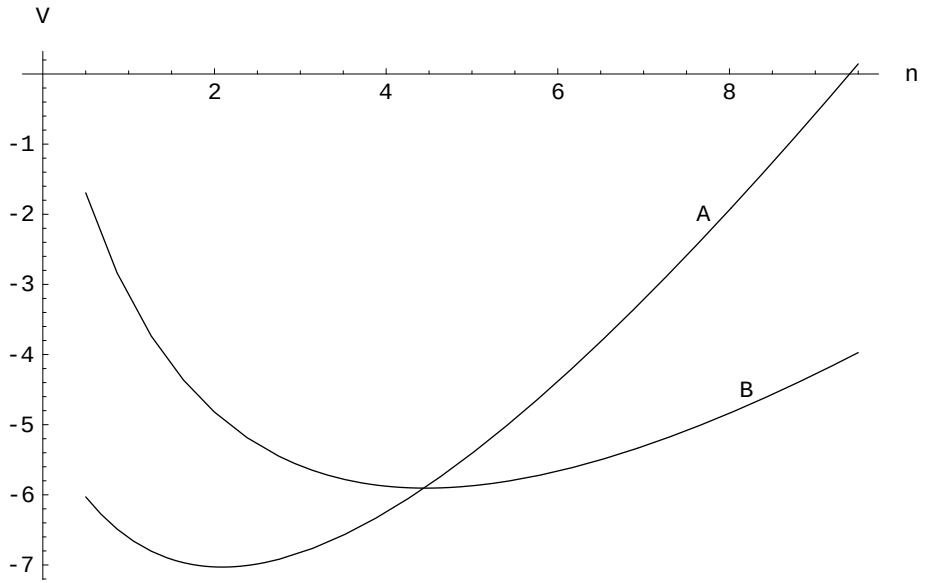

Figure 4: Two examples of potential $V(n)$.

Thus far the model is deterministic. The next step is to derive a probability distribution for n based on the potential function. With a few mild assumptions this can be done by finding a stationary solution to the so-called Fokker-Planck equation (cf. eg. Haken 1983, page 167, equation (6.110)). For the potential given above in equation (8) this results in the following equation for the probability density function $f(n)$:

$$f(n) = N \exp\left(\frac{-2(n\omega_1 - (r\omega_1 + \omega_2)\log(r+n))}{Q} \right) \qquad (9)$$

Where N is a normalizing constant and Q is the so-called diffusion coefficient which roughly stated determines how much "noise" the system contains. To finish out our example we present in Figure 5 probability density functions for the two examples of Figure 4, calculated with an arbitrary value of $Q = 1.1$.

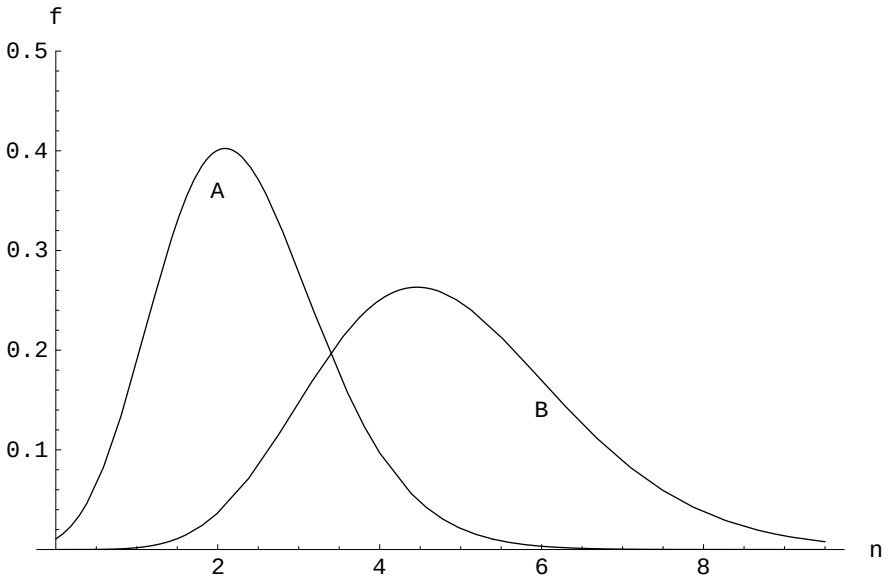

Figure 5: Probability density functions for examples in Figure 4.

Although the right hand curve (curve B) in Figure 5 superficially resembles the empirically estimated distribution shown in Figure 2, we wish to emphasize that the exact form of the theoretical distribution curve should not be taken too seriously. For one thing, it is certainly plausible that effort expended on synchronizing is not exactly proportional to $h(n)$ but rather to some transformation thereof which might radically change the shape of the potential well and the resulting probability distribution. For another thing, we have not taken into account many possibly systematic influences such as syllable type or the possibility that more than two rhythms are coupled hierarchically (but cf. Nieminen, 1996; O'Dell & Nieminen, 1998; 1999; 2001 for possible extension of the basic model to include complexities such as these), lumping all other influences into a single diffusion coefficient. Also we have not addressed the discrete nature of n.

In spite of this, one feature of the distribution is likely to be quite robust, namely the location of the mode, corresponding to the n which most closely approaches the ratio of eigenfrequencies of the stress and syllable oscillators (Ω) and therefore requires the least "synchronizing effort." Thus, providing the corpus is large enough, our analysis may afford a way to estimate Ω and make our model a bit more definite. For instance, in earlier work (O'Dell & Nieminen, 1998; 1999) we obtained an estimate for the relative coupling coefficient r based on linear regression of duration of interstress interval (ISI). For the same Finnish data considered here the regression equation obtained was ISI = $a + bn$ = 132 + 143n msec, and comparing with equation (7) it can be seen that $a/b = 0.92$ provides an estimate of r, but the eigenfrequencies of the model are not uniquely determined. However, if Ω is assumed known, then we will be in a position to estimate all the parameters of our model (ω_1, ω_2, r) simultaneously. If we assume a value of $\Omega = 4.5$, which is at least consistent with Figures 2 and 3, then given

our earlier results of regression on the same data, we obtain approximate parameter values $\omega_1 = 1.28$ Hz, $\omega_2 = 5.80$ Hz, $r = 0.92$. These are in fact the values used to compute the B curves in Figures 4 and 5. The values used for the A curves are compatible with the regression equation ISI = 207 + 118*n* msec presented by Fant & Kruckenberg (1989) for Swedish. These values assume a value of $\Omega = 2.1$ for eigenfrequency ratio, which is completely arbitrary, since we have no relevant data for Swedish.

4. CONCLUSION

It appears likely that coordination of stress and syllable rhythms in speech production does indeed have an effect on length of stress group. Of course thorough investigation of this phenomenon would require that much more data be examined, for more languages and dialects as well as more speakers, speaking rates and speaking styles within languages. It is, however, encouraging that some suggestive results were possible even with the relatively small corpus used here.

REFERENCES

AGRESTI, Alan (1990). *Categorical Data Analysis*. New York: John Wiley & Sons.

BARBOSA, Plínio A. (2000). "Syllable-timing in Brazilian Portuguese": uma crítica a Roy Major. D.E.L.T.A. (16) 2: 369–402.

_____. (2001). Generating duration from a cognitively plausible model of rhythm production. *Proceedings of the Seventh European Conference on Speech Communication and Technology (Eurospeech 2001)*. Aalborg, Denmark, September 3–7. v. 2: 967–970.

BHAT, U. Narayan (1972). *Elements of Applied Stochastic Processes*. New York: John Wiley & Sons.

ERIKSSON, Anders (1991). *Aspects of Swedish Speech Rhythm*. Department of Linguistics, University of Göteborg.

FANT, G.; KRUCKENBERG, A. (1989). Preliminaries to the study of Swedish prose reading and reading style. *Quarterly Progress and Status Report, Speech Transmission Laboratory*, KTH. 2: 1–83.

GILKS, W.R.; RICHARDSON, S.; SPIEGELHALTER, D. J. (eds.) (1996). *Markov Chain Monte Carlo in Practice*. London: Chapman and Hall.

HAKEN, Hermann (1983). *Synergetics: An Introduction*. New York: Springer-Verlag.

KARLSSON, Fred (1983). *Finnish Grammar*. Juva: WSOY.

KOPELL, Nancy (1988). Toward a theory of modelling central pattern generators. A. H. COHEN; S. ROSSIGNOL; S. GRILLNER (eds.) *Neural Control of Rhythmic Movement in Vertebrates*. New York: John Wiley & Sons. pp. 369–413.

NIEMINEN, Tommi (1996). *Suomen kielen puherytmi*. Unpublished pro gradu thesis. Department of Finnish language and general linguistics, University of Tampere, Finland.

O'DELL, Michael L. & NIEMINEN, Tommi (1998). Reasons for an underlying unity in rhythm dichotomy. *Proceedings of the Finnic Phonetics Symposium.* August 11–14, 1998. Pärnu, Estonia. Linguistica Uralica XXXIV. 3: 178–185. Available at: <http://www.uta.fi/~sktoni/tiede/odell-nieminen98.pdf>.

_____. (1999). Coupled oscillator model of speech rhythm. *Proceedings of the XIVth International Congress of Phonetic Sciences*, San Francisco. 1–7 August. v. 2. pp. 1075–1078. Available at: <http://www.uta.fi/~sktoni/tiede/odell-nieminen99.pdf>.

_____. (2001). Speech rhythms as cyclical activity. STIINA Ojala ; Jyrki TUOMAINEN (ed.): *21. Fonetiikan päivät*. Turku 4.–5.1.2001. Turun yliopiston suomalaisen ja yleisen kielitieteen laitoksen julkaisuja 67. Turku 2001. 159–165. Available at: <http://www.uta.fi/~sktoni/tiede/odell-nieminen01.pdf>.

PORT, Robert F.; VAN GELDER, Timothy (eds.) (1995). *Mind as Motion*. *Explorations in the Dynamics of Cognition.* Cambridge, Mass.:The MIT Press.

PRIGOGINE, Ilya (1997). *The End of Certainty: Time, Chaos, and the New Laws of Nature*. New York: The Free Press.

SPIEGELHALTER, D. J.; THOMAS, A; BEST, N. G. & GILKS, W. R. (1996). BUGS: Bayesian inference Using Gibbs Sampling, Version 0.5, (version ii). MRC Biostatistics Unit, Cambridge.

APPENDIX: BUGS MODEL DESCRIPTION

```
model;
{
# Q is the maximum number of syllables per stress group
   for(i in 1 : Q) {

# this transforms the log-linear parameters pi[i] into probabilities
# p[i] for stress group with i syllables
      log(phi[i]) <- pi[i]
      p[i] <- phi[i]/sum(phi[])

# calculate the probability that p[i] is the mode
      modepi[i] <- step(rank(pi[],i)-Q)
   }

# R is the total number of patterns
   for(k in 1 : R) {
      for(i in 1 : Q) {

# log-linear model of independence
# n[k,i] holds the empirical counts
         log(m[k,i]) <- alpha + pi[i] + lambda[k]
         n[k,i] ~ dpois(m[k,i])
      }
   }

# non-informative prior distributions for
# log-linear parameters
   pi[1] <- 0;
   for(i in 2 : Q) {
      pi[i] ~ dnorm( 0.0,1.0E-6)
   }
   lambda[1] <- 0;
   for(k in 2 : R) {
      lambda[k] ~ dnorm( 0.0,1.0E-6)
   }

   alpha ~ dnorm( 0.0,1.0E-6)
}
```