

Predicting Prosodic Phrasing

Pepi Stavropoulou

Supervisors: Alice Turk and Caroline Heycock



Master of Science

in

Speech and Language Processing

Department of Theoretical and Applied Linguistics

University of Edinburgh

2002

Abstract

Prosodic phrasing is an inherent part of the language mechanism. The conditions, though, that determine it, even though they have been a research topic for linguists for over two decades, are far from precisely determined and widely accepted. This thesis will present the results of a project in which an attempt was made to explore the reasons underlying the speakers' choice of prosodic phrasing and identify regularities or constraints, which have been proposed over the years to govern it.

The work departs from the comparison between two different proposals on the relation between syntactic and prosodic structure, namely Truckenbrodt's (1999) syntax-based approach and Steedman's (2000*b*) Information Structure based theory, and ends up identifying a third proposal, that of Nespors & Vogel's (1986), as the one which better predicts the partition of the test utterances into Phonological Phrases.

All proposals were assessed with respect to ToBI boundary strength tiers in a set of 451 annotated sentences of the BURadio News corpus.

Acknowledgements

Firstly, I would like to thank my two supervisors, Dr. Alice Turk and Dr. Caroline Heycock for their invaluable comments, feedback and advice on this project; particularly Alice Turk for suggesting the topic of the project to me. Also, many thanks to Dr. Simon King, the organiser of this MSc. course, who has been extremely supportive and helpful throughout this year. Furthermore, I would like to express my gratitude to the computer support staff in the department, and especially Michael Bennett for solving every single ‘Unix’ problem that came up. Last but not least, I would like to thank my course-mates, who made this past year much less difficult and much more worthwhile as well.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text. This work has not been submitted for any other degree or professional qualification except as specified.

(Pepi Stavropoulou)

Table of Contents

1 Preliminaries	1
1.1 Introduction	1
1.2 Phonological Phrases and the Prosodic Hierarchy	2
1.3 The prosodic component of the grammar	4
1.4 A constraint-based approach. H. Truckenbrodt’s proposal	7
1.5 A Combinatory Categorical Grammar Model of Prosody. M.Steedman’s proposal	12
2 Test Corpus and Method	19
2.1 Speech Style and Corpus	19
2.2 Corpus Annotation	20
3 Evaluation of Truckenbrodt’s proposal	23
3.1 General Presentation	23
3.2 Analytical Presentation	29
3.2.1 Phrase Breaks predicted incorrectly	29
3.2.2 Phrase breaks not predicted	36
4 Evaluation of M. Steedman’s proposal	53
4.1 Information Structure Identification in complex texts	53

4.2	Criteria for Identifying Information Structure	54
4.3	The choice of Information Structure domain	56
4.4	Evaluation Results	58
5	Conclusions, Applicability and Future Directions	67
A	Test Corpus	75
	Bibliography	102

List of Figures

1.1	Correspondence of prosodic constituent hierarchies from the literature	3
1.2	Prosodic constituent boundaries	3
1.3	One view of the role of the Prosodic Component in the Grammar . . .	5
1.4	Introduction of a Prosodic component in the Minimalist model	7
1.5	An example of the application of the constraints proposed by Truckenbrodt (1999)	11
1.6	An example of flexible syntactic structure in CCG	12
1.7	A CCG Model of Grammar	13
1.8	Information Structure and Contrast	16
1.9	The two dimensions of IS. Theme-Rheme and Background-Focus division.	17
3.1	The interaction between Wrap-XP and Align-XP/R constraints	24
3.2	A domination-containment distinction	25
3.3	A typical sentence in Truckenbrodt's corpus	26
3.4	A typical sentence in the BU Radio corpus	27
3.5	Occurrences of SV patterns with respect to the length of the subject .	32
3.6	ditransitive verbs	40
3.7	Restructuring cases for phrases of different length	45

3.8	Cases in which restructuring did not occur for phrases of different length	45
3.9	Coordination Structures	47
3.10	DP Structure	49
3.11	Extended projections	52

List of Tables

2.1	Correspondence between ToBI break indices and prosodic components	21
3.1	Evaluation of Truckenbrodt's proposal	28
3.2	Evaluation of Truckenbrodt's proposal(Clauses delimited with PP breaks)	28
3.3	Evaluation of Truckenbrodt's proposal, excluding sentence-final cases(Clauses delimited with PP breaks)	28
3.4	Phrase breaks predicted incorrectly (Truckenbrodt's proposal)	30
3.5	Distribution of phrase breaks at subordinate clauses' edges	30
3.6	Proper names	34
3.7	Distribution of phrase breaks before adjunct structures	35
3.8	Phrase breaks not predicted(Truckenbrodt's proposal)	36
3.9	Complement-adjunct distinction in verb phrases	37
3.10	Cases in which the distinction was easier to draw	38
3.11	Complement-adjunct distinction in Noun Phrases	38
3.12	Complement-adjunct distinction in Noun Phrases(Cases in which the distinction was easier to draw)	38
3.13	The insertion of <i>phrase head breaks</i> as a function of the <i>immediately following</i> parameter.	39
3.14	Complements and adjuncts occurring in <i>immediately following</i> the head and <i>not immediately following</i> position.	39

3.15	Minor Phrase Boundaries	42
3.16	Evaluation of Nespor and Vogel's proposal	43
3.17	Restructuring in three or more words sequences	44
3.18	Restructuring as a function of the following phrase's length	44
3.19	Coordination structures	48
4.1	Evaluation of Steedman's theory	58
5.1	Evaluation of Nespor and Vogel's proposal	69
5.2	Evaluation of Truckenbrodt's proposal	70
5.3	Evaluation of Truckenbrodt's proposal(Clauses delimited with PP breaks)	70
5.4	Evaluation of Nespor and Vogel's proposal	70
5.5	Evaluation of Nespor and Vogel's proposal(Applying restructuring) .	71

Chapter 1

Preliminaries

1.1 Introduction

This project's primary aim is to investigate the factors that determine the prosodic partition in an utterance, and particularly its correlation to syntax, discourse and length. An insight into the role of these factors would contribute to a more accurate prediction of phonological phrasing and hence more intelligible and natural synthetic speech. In order to explore their influence and relevance, two differently oriented approaches to the Syntax-Phonology interface were compared, namely

- the set of constraints developed by H.Truckenbrodt(1999) within an Optimality Theory framework and building on an edge-based approach (Selkirk 1986) to the conditions governing the syntactic and phonological component interface,
- M.Steedman's(2000) theory according to which the prosody of the utterance is determined in terms of Information Structure alone, the latter reflected on the more flexible syntactic surface structure of Combinatory Categorical Grammar,

To be more specific, the aspect of the Syntax-Phonology interface, which was explored and is presented in this paper, involves the division of the utterance into *Phonological Phrases*.

1.2 Phonological Phrases and the Prosodic Hierarchy

Phonological Phrases form a prosodic unit which has been argued to bear a close relation to Syntax and has thus been often defined in syntactic terms. To avoid any reference to Syntax at this point- whose relation to Phonology is after all the issue in question- we will simply equate the Phonological Phrase with Pierrehumbert & Beckman's (1988) Intermediate Phrase and define it in terms of a particular tune, namely a sequence of one or more pitch accents delimited by a phrase tone as opposed to an explicit boundary tone. The same assumption underlies the ToBI labelling system, which was used in the annotation of the test corpus.

As a prosodic constituent the Phonological Phrase is part of a larger system, which is hierarchically organised, namely the Prosodic Hierarchy. With respect to the latter a Phonological Phrase can be defined- continuing with Pierrehumbert & Beckman's (1988) terminology- as an intermediate level between the Full Intonational Phrase (a nuclear accent with optional pre-nuclear accents delimited by a phrase boundary) and the Clitic Group (content word with optional functional words).

Even though there is still disagreement regarding the number and the exact definition of the constituents of the Prosodic Hierarchy, a generalised mapping between different proposals and terminologies is possible, especially for the constituents higher in the hierarchy, for which there is a fairly good degree of agreement. Figures 1.1 and 1.2 (adopted from Shattuck-Hufnagel & Turk (1996)) show the correspondence between different categorisations and a parsing example respectively.

The organisation of the constituents within the hierarchy conforms to certain conditions, which we will also be assuming for the evaluation of Truckenbrodt's (1999) proposal. Following are these conditions as they appear in Selkirk (1995) :

Constraints on Prosodic Domination
(Where C_n is some prosodic category)

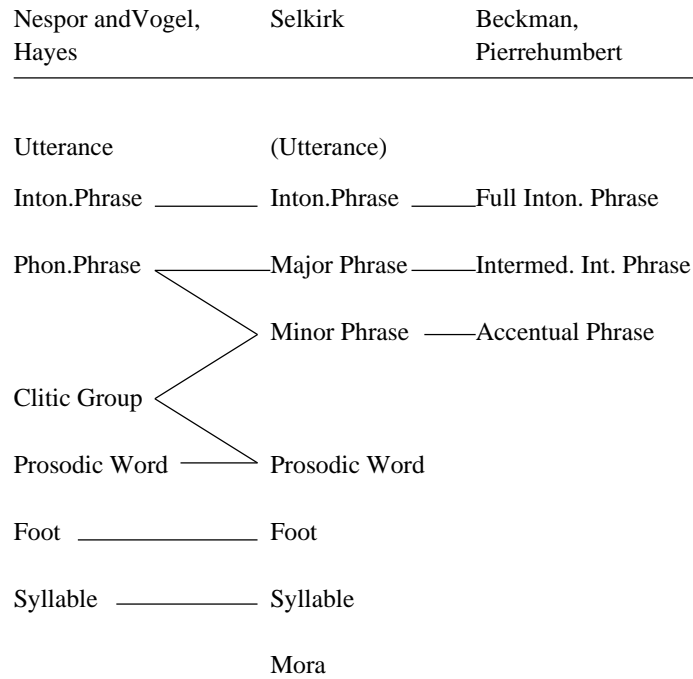


Figure 1.1: Correspondence of prosodic constituent hierarchies from the literature

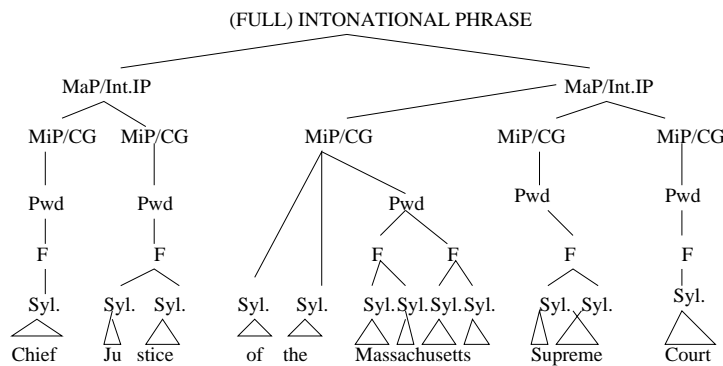


Figure 1.2: Prosodic constituent boundaries

Layerdness No C_i dominates a C_j , $j > i$

e.g. No Clitic Group dominates an Intermediate Phrase

Headedness Any C_i must dominate a C_{i-1} (except if C_i is a syllable)

e.g. A Full Intonational Phrase must dominate an Intermediate Intonational Phrase.

Exhaustivity No C_i immediately dominates a C_j , $j < i-1$

e.g. No Intermediate Phrase immediately dominates a Clitic Group.

Nonrecursivity No C_i dominates a C_j , $j=i$

e.g. No foot dominates a foot.

In several proposals on the syntax-phonology mapping- including the one we will be taking a closer look at (Truckenbrodt 1999)- the above constraints are stated as ranked and violable in terms of Optimality Theory. And while Layeredness and Headedness appear not to be violated and are thus ranked first, there are instances in which violations of Exhaustivity and Non-Recursivity have been shown to stand (see for example Hayes (1995), Inkelas (1989), Selkirk (1995), Ladd (1986), Peperkamp (1997)) , and these constraints are thus ranked relatively lower. Their ranking, however, is beyond the scope of this project and all constraints on Prosodic Domination are considered for our purposes inviolable¹.

1.3 The prosodic component of the grammar

The Prosodic Hierarchy essentially represents the organisation of the so called Prosodic Component of the model of Grammar. The postulation of such a separate component has been supported by a lot of linguists, such as Selkirk (1984), Hayes (1995), Nespor & Vogel (1986), Pierrehumbert & Beckman (1988), Inkelas (1989), Ladd (1996), Truckenbrodt (1999), and many others, who recognize the complexity of the prosodic representation and the need for an independent level, where factors diverse in nature

¹Note, however, that Truckenbrodt (1999) resorts to the violability of a ‘Non-recursivity’ constraint, in order to explain certain cases of prosodic phrasing in his data. Ladd (’86, ’92, ’96) proposes the existence of nested Intonational Phrases to account for subtle differences in boundary strength and further suggests the possibility of even eliminating the distinction between Full Intonational Phrases and Intermediate Intonational Phrases and use the notion of Compound Intonational Phrases instead

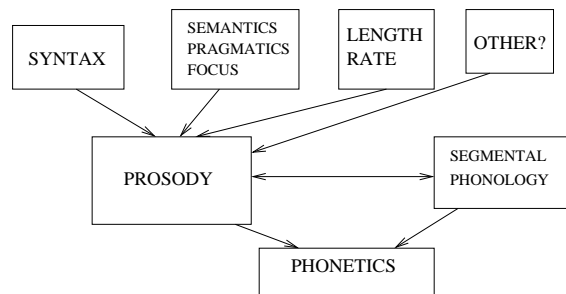


Figure 1.3: One view of the role of the Prosodic Component in the Grammar

can be integrated.

The justification for each prosodic constituent lies mainly on the fact that they have been shown to serve as the domain of application of different phonological rules. And that domain, on the one hand, is quite often anisomorphic to any of the syntactic constituents, while on the other hand, seems to be sensitive to factors related not only to syntax, but also to semantics, discourse, speech rate, phrase length and balance (Figure 1.3. One view of the role of the Prosodic Component in the Grammar (Shattuck-Hufnagel & Turk 1996)).

That taken into consideration, a theory which claims there can be a straightforward account for phonological phenomena based on their relation to syntax alone, and phonological rules apply to syntactic structure directly, seems rather simplistic. Even for units such as the Phonological Phrase, which are argued to be greatly dependent on the syntactic representation, the postulation of such an independent prosodic component is argued to be theoretically simpler than identifying subsets of syntactic constituents to serve as the domain of phonological rules application instead (Nespor & Vogel 1986). Furthermore, it is argued to capture generalisations that would have otherwise been missed.

Finally, the Prosodic Component seems to be perceptually motivated as well, as it forms- according to some theorists at least- the first stage in speech perception see Nespor & Vogel (1986).

Nevertheless, the postulation of such a component does not come without problems. Given the branching T configuration of the Minimalist Model (see figure 1.4), which many of the theories supporting an individual Prosodic Component assume, there can be no interaction between the covert component² and Phonological component³ after the point of Spell-Out⁴, at which the derivation splits up and elements interpretable at the LF interface⁵ are left behind for the covert component, while elements interpretable at the PF interface⁶ are spelt out, carried over to the Phonological component, and receive an overt form (pronunciation). The conditions, however, on the derivation of the Prosodic Structure require an interface with the semantic component as well (see figure 1.4).

Therefore, in order for the necessary semantic information to be accessible by the Prosodic Component and still maintain the basic configuration and assumptions for the theoretical model, one would have to assume that all this information is available pre Spell-Out and is both left behind for λ and carried over to π . Accordingly, there have been proposals for relevant annotated syntactic structures such as Selkirk's (1984) intonated surface structure or the F(ocus)-marked constituents in Syntax (Jackendoff 1972).

In any case, though -whether the necessary information is available pre Spell-Out or whether the prosodic and the LF representations communicate- the model would be considerably further complicated in contrast to the general tendency within Minimalism for simpler representations and abandonment of unnecessary levels (Apart from

²The subsystem that continues the computation to LF after Spell-Out mapping a linguistic expression to its logical form (Chomsky 1995).

³Phonological component: The subsystem that maps a linguistic expression to its phonetic form (Chomsky 1995).

⁴Spell-Out is the point in the computation of a linguistic expression at which the derivation splits into two heading towards the two interface levels, LF and PF. It can be said to correspond roughly to the Surface Structure in the previous models

⁵The level of representation 'that interfaces with a speaker's general knowledge and with extralinguistic cognitive systems' (Marantz 1995)

⁶The level of representation 'that interfaces with the perceptual system in speech recognition and with the articulatory system in speech production' (Marantz 1995). Informally, the PF representation includes phonetic features giving instructions on how a sentence is pronounced, while the LF representation includes semantic features describing a sentence's meaning.

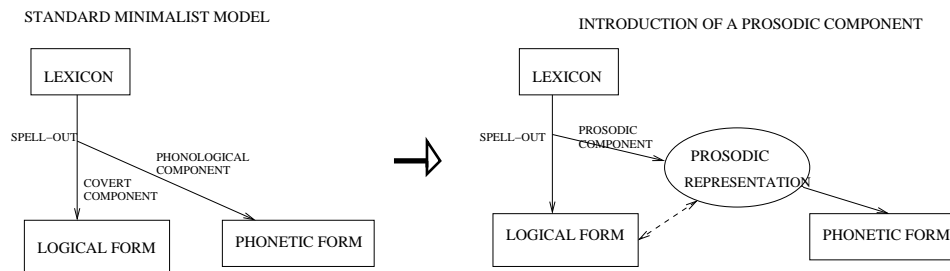


Figure 1.4: Standard Minimalist Model: Language is a computational system which, given an array of lexical items, constructs a linguistic expression, that is a pair (π, λ) , where π is a PF representation and λ is an LF representation. At some point (Spell-Out point) of the computation the abstract linguistic structure which has been formed so far, is stripped off the phonetic features, which are then processed by the phonological component and mapped to PF, while semantic features are left behind for the covert component, to be interpreted at LF. After the point of Spell-Out, there is no further interaction between the PF and LF components, and any computation on the road to LF does not affect the pronunciation of the sentence. Apparently, there are also no other levels -except for PF and LF- of representation.

The postulation of an independent prosodic component, on the other hand, requires another level of representation, which would interact with LF as well, in order to access the necessary semantic information that has been shown to affect prosody.

PF and LF there, no other levels of linguistic structure are acknowledged).

1.4 A constraint-based approach. H. Truckenbrodt's proposal

The first proposal that will serve as the starting point for 'exploring' the syntax-phonology mapping is the set of constraints suggested by H. Truckenbrodt in his 1999 paper. Truckenbrodt is essentially building on an edge-based approach to the syntax-phonology interface formulated initially by Selkirk (1986). According to this approach, the left or

right edge of a particular syntactic constituent aligns with the edge of the corresponding constituent of the Prosodic Hierarchy⁷. This mapping constraint was afterwards generalised (McCarthy & Prince 1993) to include different types of alignment. Formally ((Selkirk 1995):

Generalised Alignment

Align(α Cat, E; β Cat, E)

- Align (GCat, E; PCat, E)
- Align (PCat, E; GCat, E)
- Align (PCat, E; PCat, E)

(GCat ranges over morphological and syntactic categories;
PCat ranges over prosodic categories; E= Right or Left)

In accordance with these conditions the edge of a syntactic maximal projection is argued to coincide with the edge of a Prosodic Phrase.

Truckenbrodt (1999) follows Selkirk also in stating the conditions on the Syntax-Phonology interface as constraints in optimality theoretic terms. According to Optimality Theory, there is a universal set of constraints on ‘competing’ representations. The importance of each constraint is represented through its ranking within a constraint hierarchy. The ‘winning’, i.e. the grammatical representation, is the optimal one, that is the one that violates the fewest or/and least important constraints. Finally, typological differences across languages are due to the different ranking of each constraint within the hierarchy.

⁷The end-based relation between syntactic and phonological structures was originally observed by Chen (1985)

(Truckenbrodt 1999) adds to an existing set of constraints what he calls the Wrap-XP constraint, which is claimed to be a better alternative to the lexical government condition proposed by Hale & Selkirk (1987). According to the latter, maximal projections which are lexically governed do not trigger a Phonological Phrase boundary. Wrap-XP states instead that each maximal projection must be contained in only one Phonological Phrase. The effect of both constraints is more or less similar. However, Wrap-XP was shown to achieve better coverage across the languages tested in Truckenbrodt (1999). Furthermore the importance of lexical government seems to decrease in the latest stages of Generative Grammar⁸ Still, in both cases, the Prosodic Component requires access to more complex syntactic notions such as domination and containment rather than merely edges, as it had been earlier suggested (Selkirk 1986).

The set of constraints we will be assuming for evaluating Truckenbrodt's (1999) proposal is the following (along with the general constraints on prosodic domination):

Align-XP/R For each XP there is a PP such that the right edge of the XP coincides with the right edge of the PP, where XP is a maximal projection and PP is a Phonological Phrase. This constraint represents the end based mapping assumption for Major Phonological Phrases in English, whose right end is supposed to align with the right end of Maximal Projections.

Wrap-XP Each XP is contained in one Phonological Phrase. Align-XP/R and Wrap-XP are often contradictory, since Align-XP/R favours the postulation of phonological boundaries, while Wrap-XP prohibits it. Take for example the following syntactic configuration corresponding to a ditransitive verb and its complements, where VP2 and VP3 are the complements included in VP1.

(V1 VP2 VP3)_{VP1}

Wrap-XP would rule out the following prosodic partition generated in accor-

⁸In previous versions of the theory, case assignment, for example, was conditioned on the relation between the head and its lexically governed complement. Within the Minimalist Program, however, case feature checking is only possible in a head-specifier relation. In general, the relation of government no longer allows one element to license another.

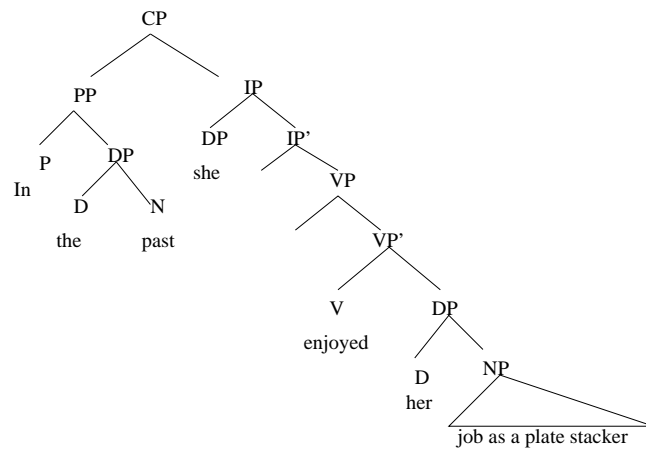
dance with the Align-XP constraint

(V1 VP2)(VP3)

and impose a partition in which VP1 is wrapped in a single phonological phrase (V1 VP2 VP3).

Lexical Category Condition(LCC) Constraints relating syntactic and prosodic categories apply to lexical syntactic elements and their projections (NP,VP, AP), but not to functional elements and their projections (CP, IP, DP), or to empty (phonetically null) syntactic elements and their projections. Thus in the example in figure 1.5 no phonological phrase break is inserted after the subject ‘she’, since pronouns have the status of function words. A phrase break would be triggered with respect to the projection of a noun, a verb or an adjective. As far as prepositions are concerned, being a category which stands somewhere between functional and lexical elements, they will for our purposes be considered not to form a lexical category, since -with respect to Phonology at least- they seem to behave similarly to functional rather than lexical heads (see Nespor & Vogel (1986)).

***P-Phrase** the insertion of phonological phrases is penalised altogether. This constraint is often used to inactivate others which are consequently ranked below it.



(In the past) (she enjoyed her job as a plate stacker)

Figure 1.5: An example of Phonological Phrase partition (denoted by brackets) with respect to the proposed constraints. According to Align-XP/R a phrase break is triggered at the right edge of the NP *past*, the VP *enjoyed...plate stacker* and the NP *job...stacker* (the last two phrase breaks coincide at the end of the sentence). Projections of functional words such as *she* do not trigger a phrase break because of the LCC condition, and they are included in the phonological phrase of the lexical maximal projection they are associated with. Finally, Wrap-XP is not violated since all lexical XPs are included in one Phonological Phrase.

1.5 A Combinatory Categorical Grammar Model of Prosody.

M.Steedman's proposal

In a different approach Steedman (2000*b*) proposes a fairly straightforward account for the prosodic structure according to which the latter is a function of Information Structure alone. The need for an individual prosodic component in the grammar is thus rejected, as the prosody of the utterance is claimed to be determined in terms of the utterance's information organisation, which in turn coincides with the syntactic structure. In short, instead of including an additional prosodic component in the grammar, both the syntactic and the prosodic levels are merged in a single level reflecting the information structure within the utterance. Such direct correspondence between different structural levels is possible given the flexible syntactic structure of CCG(Combinatory Categorical Grammar), which allows for both traditional and non traditional constituents, which arise in prosodic phrasing, to be derived (see figure 1.6).

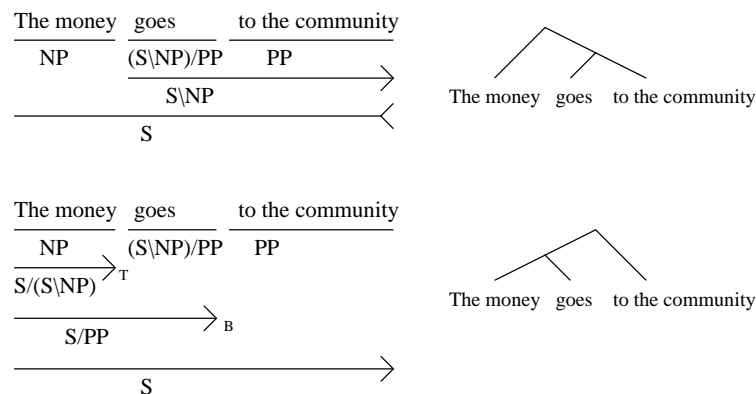
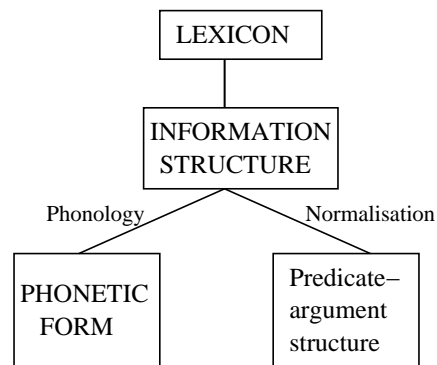


Figure 1.6: An example of flexible syntactic structure in CCG. Each derivation results in a difference in constituency. In the first case, the traditional one, the verb first combines with its complement to form a VP and then the VP combines with the subject. In the second case, however, the verb first combines with the subject and then the new constituent combines with the verb's complement. The choice between the two alternative representations is determined by the Information Structure of the sentence, reflecting a theme-rheme distinction.

The result is the stipulation of a simple -and indeed attractive in its simplicity- model of the language mechanism (figure 1.7).



[Steedman, 2000]

Figure 1.7: A CCG Model of Grammar. Intonation and Syntactic structure are subsumed under Information Structure. Information Structure is, thus, the only explicit level of representation among the three, constituting -along with the predicate-argument structure- the LF level. In order to get from the structural representation of IS to the canonical predicate-argument structure, normalisation occurs as θ - ρ marking 'disappears' affecting only the database (by identifying θ - ρ sets).

As a consequence of such flexibility a simple even sentence will have multiple derivations, which would correspond to the same predicate-argument structure but -according to Steedman (2000*b*)- to different information structure as well. Steedman's (2000*b*) claim is that each derivation corresponds to a different organisation of the information in the utterance and that by integrating intonation structure (as a marker of IS) into syntax (as a representation of IS) one can uniquely identify (or at least narrow down the possibilities of) a single derivation as the correct one (the correct one being the one representing the speaker's actual information organisation for the given utterance in the particular context).

To be more specific, different pitch accents are claimed to function as either theme or rheme markers. Adopting Pierrehumbert & Hirschberg's (1990) description of intonation contours, Steedman (2000*b*) identifies tones L*+H and L+H* within their system as markers of theme, and tones H*, L*, H*+L and H+L* as markers of rheme⁹, the theme being *the part of the utterance which connects it to the previous discourse* reflecting the context and rheme being *the part of the utterance which advances the discourse* affecting the context¹⁰. Boundaries, on the other hand, do not actually determine discourse function, i.e. 'themehood' or 'rhemehood', but rather delimit the theme's or rheme's projection from single words, and mark the theme or rheme part as being the hearer's or the speaker's. In particular each pitch accent (marking either theme or rheme) associated with a particular word, sets the value of that word's INFORMATION feature (otherwise unspecified) to theme or rheme accordingly. 'Themehood' and 'rhemehood' is then projected from individual words to phrases through unification. The range of the projection and thus the derived constituent is constrained by the presence of phonological boundaries (phrase or boundary tones).

At the end, the partition of the utterance into Phonological Phrases should coincide with the partition of the utterance into theme and rheme.

The process described above is in line with -or more precisely a result of - a two-dimensional view of Information Structure. According to that view the theme-rheme partition constitutes the first dimension of IS. Theme and rheme are both members of a set of alternative themes and rhemes respectively, determined by the particular context. The theme alternative set is formalised as a set of functional abstractions, while the rheme alternative set consists of the possible instantiations of the available alterna-

⁹The association of different intonation tunes with different discourse meanings have often been suggested, even though specifying exactly what that association is and what the particular aspects of that meaning are, has been rather difficult and controversial (see among others Pierrehumbert & Hirschberg (1990), Selkirk (1984), Ladd (1996), Gussenhoven (1983), Halliday (1967), Jackendoff (1972), Pierrehumbert & Hirschberg (1990))

¹⁰These are examples of theme rheme partition corresponding to the different surface structure representations in figure 1.6.

Q: What goes to the community?

A: (The money)_ρ(goes to the community)_θ

Q: Where does the money go?

A: (The money goes)_θ(to the community)_ρ.

For a more detailed discussion on the properties of theme and rheme, and Information Structure Identification see Chapter 4.

tive themes.

For example in the following discourse:

Q: Aretha and Bertha are both Clyde's girlfriends. But who does Clyde love?

A: Clyde likes Bertha

the theme alternative set established by the question would be

$\lambda x.\text{love}'(x, \text{clyde}')$

$\lambda x.\text{like}'(x, \text{clyde}')$...

while the rheme alternative set would consist of all the possible instantiations

$\text{love}'(\text{bertha}', \text{clyde}')$

$\text{love}'(\text{aretha}', \text{clyde}')$

$\text{like}'(\text{bertha}', \text{clyde}')$

$\text{like}'(\text{aretha}', \text{clyde}')$

$\text{like}'(\text{julie}', \text{clyde}')$...

And the final saturated proposition with respect to the response would be

$\text{like}'(\text{bertha}', \text{clyde}')$

The second dimension of Information Structure involves a focus-background contrast within theme and rheme separately associated with individual words. Contrast is signalled by the focus constituent, which is related to the possibility of different referents made available by the discourse and the situational context (see figure 1.8 (adopted from Komagata (1998)), and is marked by a pitch accent as opposed to background material which remains unmarked by pitch accents or edge tones¹¹.

¹¹It seems that -from a Phonological point of view- Steedman's approach is primarily influenced by a focus-to-accent view (focused words are accented and accented words must mean that they are focused). In his paper *Syntax-Phonology Interface*, though, he recognizes the possibility of non-contrastive secondary accents and separates them from primary accents, which he relates to focus.

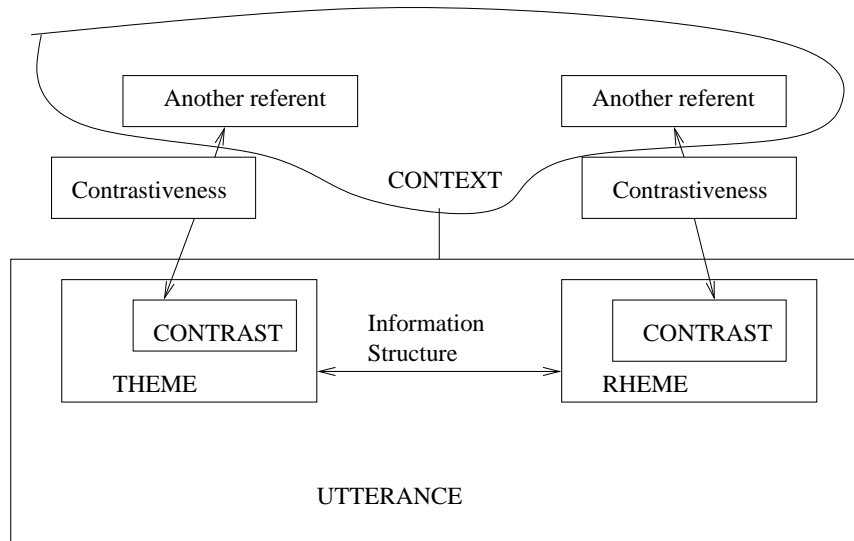


Figure 1.8: Information Structure and Contrast. Focused words within theme or rheme are in a contrastive relation with other referents within the discourse or situational context. In the previous discourse on page 15 for example, *bertha'* (focus within rheme coinciding here with the whole ρ -part of the utterance) was contrasted with other referents such as *aretha'*, *julie'* etc. (referents may not be exhaustively known to the hearer), that were made available by the context.

Focused words then ‘contribute to distinguishing the theme and the rheme of an utterance from other alternatives made available by the context’ (Steedman 2000*b*). That is the choice of focus within the theme restricts the theme alternative set selecting one theme-proposition and going on to generate the new rheme alternative set, while the choice of focus within the rheme-part restricts the rheme, which in turn reduces the theme to form the final saturated proposition out of the alternative set of final functional abstractions. For example consider the following discourse (figure 1.9, from Steedman (2000*a*))

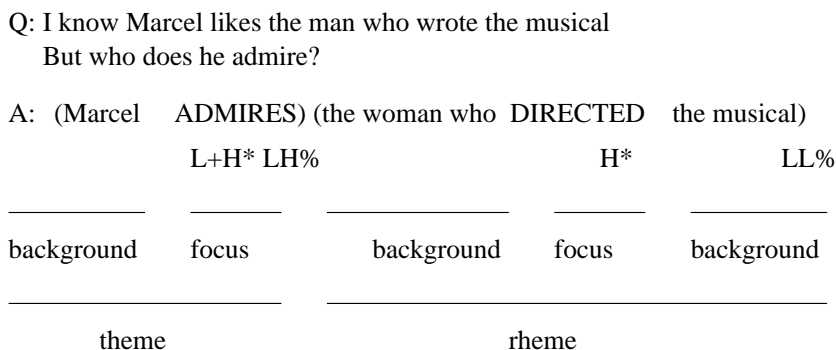


Figure 1.9: The two dimensions of IS. Theme-Rheme and Background-Focus division.

The theme alternative set would be:

$\lambda x.$ admires’ marcel’

$\lambda x.$ likes’ marcel’

The focus within the theme would then restrict the theme alternative set to

$\lambda x.$ admires’ marcel’

and establish the new rheme alternative set:

admires’woman1 marcel’

admires’woman2 marcel’

The focus on *directed* -directing being the property which distinguishes the referent uniquely- finally restricts the set to one proposition.

To sum up, in predicting phonological phrasing with respect to this theory, one has to identify the theme and rheme parts of the utterance. Syntactic structure and Syntactic boundaries do coincide -as was the case with Truckenbrodt's (1999) theory- with phonological boundaries, only this time syntax is essentially a direct function of Information Structure, while the model's structural flexibility allows the derivation of all different kinds of constituents. Finally, no other factor is said to affect Phonological Phrasing.

Chapter 2

Test Corpus and Method

2.1 Speech Style and Corpus

The two proposals were tested on a set of 30 news stories produced by speaker 2fb of the Boston University Radio News Corpus. The whole corpus amounts to 451 sentences. The speaker is a professional female radio announcer and is categorised as type B. Type B speakers ‘normally pre-record and edit their stories. The stories read by type-B announcers are typically longer, more in-depth feature stories than the short news updates read by the type-A speakers. Type-B speakers are also more familiar with the material, since they wrote it’(Ostendorf, Price & Shattuck-Hufnagel 1995). Type-B speakers are thus likely to be less prone to hesitation and false parsing mistakes. In general professional announcers tend to use clear and consistent prosodic marking, while this broadcasting speech style is argued to be ‘a natural but controlled speech style, combining some of the advantages of read speech(greater control...), with some of the advantages of non-read speech (it is an accepted, naturally-occurring form of read speech)’(Ostendorf et al. 1995).

2.2 Corpus Annotation

Prosodic labelling of the corpus was performed using the ToBI annotation system. The ToBI system is based on Pierrehumbert's proposed description of intonational tunes within Autosegmental Metrical Theory, based in particular on a more simplified version of the original proposal (Pierrehumbert 1980) presented in Silverman, Beckman, Pitrelli, Ostendorf, Wightman, Price, Pierrehumbert & Hirschberg (1992). According to Pierrehumbert's analysis, intonation contours are decomposed into tonal events associated with specific time points in the utterance. The events are of two types, namely edge tones and pitch accents. Edge tones are further divided into phrase tones(H_- , L_-) and boundary tones($H\%$, $L\%$). There are two main types of pitch accents, H and L, and further combinations of these two tones. ToBI allows for the following five different types of accents:

H^* , L^* , L^*+H , $L+H^*$, $H+!H^*$,

where the * denotes the central tone, which is linked to the syllable bearing the main word stress, while ! marks a downstepped pitch accent.

ToBI further accommodates a break index tier as well. The break index system is used to describe the perceived strength of each word boundary, *the degree of juncture between each pair of words*. There are five different indices representing degrees of strength and ranging from 0 (weaker boundary) to 4(stronger boundary). In particular:

BI 0 cases marked with phonetic modifications such as palatization.

BI 1 normal word boundary.

BI 2 it is normally used for a pause associated with a group of words, which is not however accompanied by an edge tone.

BI 3 Intermediate Phrase Boundary marked by a phrase tone.

BI 4 Full Intonation Phrase Boundary marked by a phrase tone followed by a boundary tone.

BI 4	->	Full Intonational Phrase
BI 3	->	Intermediate Intonational Phrase
BI 2	->	Clitic Group
BI 1/0	->	Prosodic Word

Table 2.1: Correspondence between ToBI break indices and prosodic components

Table 2.1 shows the correspondence between the break indices and the constituents of the prosodic hierarchy.

As it can be seen from break indices 3 and 4, tonal sequences and break levels need to coordinate, something which in practice poses a problem, as Intermediate Phrase tunes are sometimes found not being accompanied by the corresponding perceived boundary (BI 3). In that case BI 2 is used instead. This is actually an instance of conflicting criteria (see Ladd (1996)) on the definition of Intermediate Phrases, the one criterion being a characteristic phrase tune and the other being perceived boundary strength.

For the purposes of this project the latter criterion (perceived boundary strength) is going to be used in the evaluation of each theory's predictions. In particular we will be focusing on break indices 3 and 4 corresponding to an Intermediate Phrase boundary and a Full Intonational Phrase boundary¹ respectively.

With respect to the reliability of the criterion the following should be pointed out:

- The level of agreement among annotators is fairly satisfactory and certainly higher than the corresponding one for tone tiers. In a labeller consistency study (Ostendorf et al. 1995), the interannotator agreement for break indices was 95% within uncertainty level², whilst for pitch accents the interannotator agreement

¹and thus -in accordance with the nature of the Prosodic Hierarchy- to an Intermediate Phrase boundary as well

²Neighbouring classes are merged, wherever there is uncertainty. ToBI uses a diacritic '·' to indicate

was initially 60%, and 81% when H* and L+H* tones were grouped together. Finally, edge tone agreement was 93% and 91% for boundary tones and phrase tones respectively.

- One might argue that annotators are influenced by Syntax in their prosodic judgments. Supporting the reliability of perceived boundary strength as an indicator of prosodic structure and the lack of any syntactic influence in labelling judgments Collier, de Pijper & Sanderman's (1993) experiments showed a good rating agreement even for delexicalized, unintelligible speech, where prosodic structure would not affect perception. Furthermore, boundary strength level distinctions have also been shown to correspond to different grades of pre-boundary segmental durations matching the different constituents in the Prosodic Hierarchy (Wightman, Shattuck-Hufnagel, Ostendorf & Price 1992).

transcriber uncertainty. For example PBI 3- indicates uncertainty between PBI 3 and 2. Similarly a PBI 4- might correspond to a PBI 3 or 4. Thus annotators' markings 3 and 4- were merged as 3, to calculate the level of interannotator agreement.

Chapter 3

Evaluation of Truckenbrodt's proposal

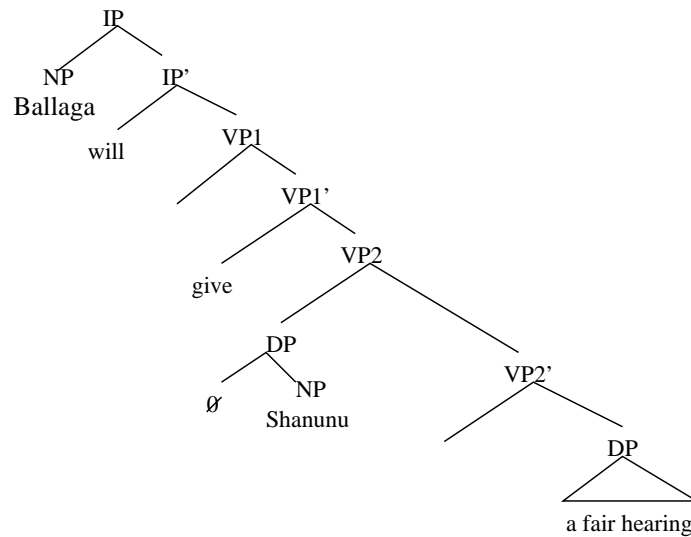
3.1 General Presentation

Our discussion begins with the evaluation of H. Truckenbrodt's proposal. Before presenting the results, the following should be made clear to the reader:

Following Truckenbrodt (1999), the Phrase Structure upon which the constraints apply, is the one developed within Minimalism. Even though X-bar entities are banned in Minimalism, they will, nevertheless, be used here as an informal representation of the sentence's structure for expository reasons.

Since the conditions on the Prosodic Interface are expressed in terms of Optimality theory, the output form is sensitive to the constraints' ranking. For the moment it will be assumed that Wrap-XP is ranked higher than Align-XP/R and we will be discussing that assumption in a later section. Figure 3.1 illustrates how exactly the two constraints interact.

Furthermore, it should be pointed out that all constraints on the syntax-phonology mapping are claimed to apply on syntactic categories rather than syntactic segments. Therefore, a distinction between the notion of domination and containment, and con-

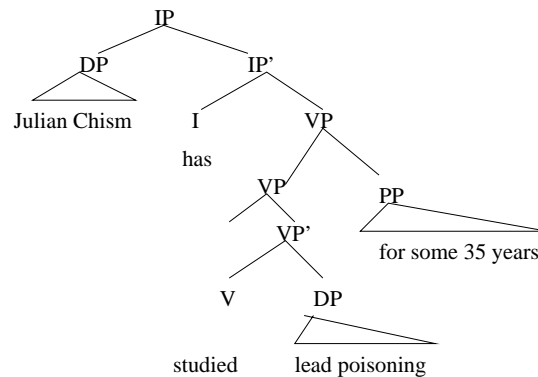


(Ballaga)(will give Shanunu)(a fair hearing)
 (Ballaga)(will give Shanunu a fair hearing)

Figure 3.1: Wrap-XP requires that all lexical XPs (thus all NPs and VPs in our example) are included in only one phonological phrase. Therefore the verb phrase ‘give Shanunu a fair hearing’ should form one phonological phrase, as in [(give Shanunu a fair hearing)]_{VP}. Align-XP/R on the other hand requires the insertion of a phrase break after every lexical XP, and hence after the NP Shanunu as well, resulting in the VP getting broken down into two Phonological Phrases. Thus, if Align-XP/R is ranked higher, Wrap-XP gets violated and the first phrasing at the bottom of the figure (NP)[(NP)(NP)]_{VP} is the one predicted. Otherwise the second phrasing (NP)[(NP NP)]_{VP} is predicted.

sequently between complements and adjuncts turns out to be rather important for the proposal's evaluation. To be more specific, while an XP node in a syntactic representation corresponds to a syntactic segment, an XP category is an abstract entity, which consists of the relevant set of syntactic segments. Further, for a constituent to be dominated by that category, that constituent requires to be dominated by every segment in that category. Since adjuncts merge with an XP node to create another XP node of the

same type, the adjunct structure is not dominated but merely contained in the maximal projection. To give an example, in figure 3.2 the prepositional phrase *for some 35 years* is an adjunct which adjoins to the VP *studied lead poisoning* to create another VP segment, and thus a two-segment VP category. So, *for some 35 years* is contained in the VP, since it is dominated only by the highest VP node, while the verb complement *lead poisoning* is dominated by the VP category, since it is dominated by both VP segments-nodes. The Align-XP/R constraint will apply only to the constituents dominated by the VP category, that is the verb *studied* and its complement *lead poisoning*, requiring the triggering of a phrase break at the right edge of the latter. Furthermore Wrap-XP is not violated, since all constituents dominated by the VP category are included in one Phonological Phrase. If on the other hand containment was chosen over domination, the predicted prosodic partition for the example would be (Julian Chism)(has studied lead poisoning for some 35 years) which turns out to give far worse results for the fine prosodic partitions in our data. In short, the XPs in the definition of each proposed constraint should be interpreted as the set of constituents dominated by the XP category, i.e. by all XP nodes.



(Julian Chism)(has studied lead poisoning)(for some 35 years)

Figure 3.2: A domination-containment distinction

With respect to the complement-adjunct distinction that is needed to be drawn, it should be said that such a distinction is not at all straightforward. The basic crite-

tion that was used was the obligatory or not presence of the constituent in question, following a more syntactically oriented approach, as obligatory presence may differ with respect to syntax, semantics or even the specific communication requirements.

Finally a problem for the proposal and its evaluation came up, because of the difference in the nature of the corpus Truckenbrodt (1999) based his proposal upon, and the test corpus used for this project. Even though it is not made explicit in Truckenbrodt (1999), it would probably be correct to assume that the utterances reflect a colloquial, normal rate speech style in contrast to the formal, broadcasting style of the BURadio News Corpus. As a result the complexity of the utterances differs greatly. Thus, a typical example of Truckenbrodt's (1999) utterances would be:

The cowboy branded the calf

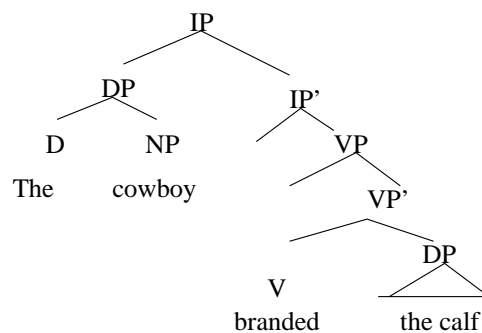


Figure 3.3: A typical sentence in Truckenbrodt's corpus

while a typical sentence in the BU corpus is the following:

The initiative process worries Beard, who says it can offer simplistic solutions for complex problems such as taxation and nuclear energy.

If Wrap-XP was active in the second utterance we would expect the circled VP (figure 3.4) to form one very big prosodic unit. In the actual case, however, the VP is broken down into seven Phonological Phrases.

It is very possible that this style difference actually matters in that one would be likely

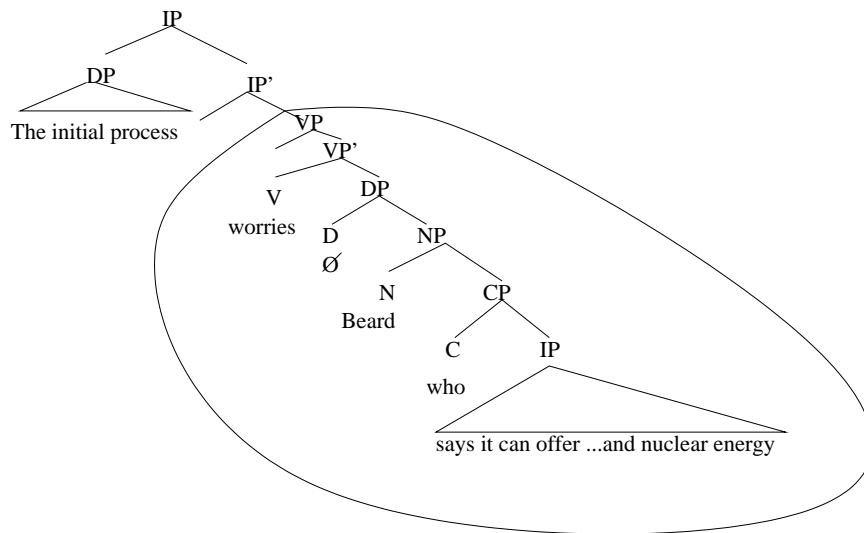


Figure 3.4: A typical sentence in the BU Radio corpus

to expect either the insertion of additional phrase breaks or phrase breaks appearing stronger in the latter case. Professional announcers have been shown to often use a much finer phonological partition, probably finer than usual (see for example Fujisaki, Hirose, Takahachi & Morikawa (1986), Price, Ostendorf, Shattuck-Hufnagel & Veilleux (1988)). Whatever the case might be though, what is for certain is that the complexity of the utterances in the two corpora poses a difficulty from the very beginning affecting the validity of the proposed set of constraints.

To overcome the problem of complexity, the proposal was also evaluated assuming that every clause is delimited by a phrase break regardless of the theory's stipulations. A further motivation for our assumption was the fact that 85% (414/487) of subordinate clauses in the corpus are separated from their superordinate clauses and the rest of the sentence with a Phonological Phrase break. Table 3.2 shows the results, given that assumption, and table 3.1 shows the results conforming exactly to the theory's predictions.

PBs NOT PREDICTED	PBs PREDICTED CORRECTLY	PBs PREDICTED INCORRECTLY
1399	1271	237

Table 3.1: Evaluation of Truckenbrodt's proposal(Clauses are not dealt with individually)

PBs NOT PREDICTED	PBs PREDICTED CORRECTLY	PBs PREDICTED INCORRECTLY
650	2020	375

Table 3.2: Evaluation of Truckenbrodt's proposal(Clauses delimited with PP breaks)

PBs NOT PREDICTED	PBs PREDICTED CORRECTLY	PBs PREDICTED INCORRECTLY
650	1569	375

Table 3.3: Evaluation of Truckenbrodt's proposal(Clauses delimited with PP breaks), excluding predicted sentence final phrase breaks

As it can be seen from the tables, the predictions that the theory makes are much more accurate when dealing with each clause separately, meaning that Wrap-XP should probably be changed to be active only clause internally. The great difference in the theory's predictions in each case, is due to the fact that even when no PP break occurs at the very beginning of the clause, there are PP breaks within that clause, which the theory would have not predicted because of the Wrap-XP constraint. In the following sentence for example there is no phrase break between the main clause verb and its sentential complement (the italicised clause), but there are several phonological boundaries within that sentential complement.

...(Dr. Ed Beard)(says *the push*) (for *do it yourself lawmaking*)(comes from voters)(who feel) (frustrated by their elected officials).

Wrap-Xp would again indicate the following phonological partition:

(Dr. Ed Beard)(says *the push for do it yourself lawmaking comes from voters who feel frustrated by their elected officials*)

Thus comparatively, even for the 15% of cases, in which no phrase break occurs at the left edge of the clause, changing Wrap-XP so as to apply within each clause separately would still yield better predictions.

3.2 Analytical Presentation

Following is a more detailed discussion of each type of data for which the proposal failed to make a correct prediction. The types of incorrect predictions fall under two broad categories, the first one being phonological boundaries which were predicted by Truckenbrodt's (1999) proposal, but did not appear in the corpus (375 PBs), and the second one being phonological boundaries, which did appear in the corpus, but had not been predicted (650PBs).

3.2.1 Phrase Breaks predicted incorrectly

Table 3.4 illustrates the types of phrase breaks that were predicted by the theory¹, but did not actually occur in the data. The following are typical examples for each case starting from the top of the table. The underlined part corresponds to the point of the utterance in which a phrase break was predicted between the two words but did not occur :

(Betty Zisk) (found that 88% percent of the battles)(were won)(by the side)(that spent the most money)

(The state will save)(about one million dollars)

(The contract)(was unveiled last May)

Under the *others* category are included trivial cases such as genitive structures

(Frank'smore recent admission)²

and phrase breaks within coordination structures

(231 men)(women and children).

We will be discussing each type of incorrect prediction separately in the next sections.

¹From now on, by theory we will be referring to Truckenbrodt's (1999) proposal adjusted so that each clause is processed separately.

²((Frank)_{NP}'s_{SD}(more recent admission)_{NP})_{DP}.

CLAUSES	72	19.8%
(SV) PATTERNS	187	49.86%
(XP-Adjunct) PATTERNS	78	20.8%
OTHERS	38	10.13%
TOTAL	375	

Table 3.4: Phrase breaks predicted incorrectly.

3.2.1.1 Clauses

As it was mentioned at the end of the previous section, in 85% of subordinate clauses, the clause was delimited by phrase breaks. The rest 15% consists of phrase breaks that were predicted incorrectly. Table 3.5 demonstrates the distribution of phrase breaks among subordinate clauses. Since the right edge of the clause was always-without any exception-aligned with a PP boundary, only left edges were considered for this table³.

	PP Break Index 4	PBI 3	PBI 2	PBI 1/0
CLAUSE TYPE				
<i>SENTENTIAL COMPLEMENTS</i>	136 (46.6%)	89 (30.5%)	34 (11.6%)	33 (11.3%)
<i>RESTRICTIVE RELATIVE CL.</i>	26 (61.9%)	13 (30.9%)	3 (7.1%)	
<i>NON-RESTRICTIVE RELAT. CL.</i>	34 (94.4%)	2 (5.6%)		
<i>ADVERBIAL CLAUSES</i>	69 (92%)	4 (5.3%)	1 (1.3%)	1 (1.3%)
<i>-ED/-ING PARTICIPLE STRUCT.</i>	36 (85.7%)	5(11.9%)		1 (2.4%)
SUBTOTAL(sent. compl. excluded)	165 (84.6%)	24 (12.3%)	4 (2.05%)	2 (1%)
TOTAL	301 (61.8%)	113 (23.9%)	38 (7.8%)	35 (7.2%)

Table 3.5: Distribution of phrase breaks at subordinate clauses' edges

First of all, non-restrictive relative clauses⁴, which are known to form their own Into-

³Unless clauses were in a sentence initial position

⁴e.g. (David Skondras)_{FIP}(*who is gay*)_{FIP}(is plenty angry)(but not at Barney Frank,)_{FIP}(*who paid for sex*)_{FIP}...

national Phrase (e.g. see Selkirk (1978)), are in all but two cases (in which there is an Intermediate Phrase break instead) produced as such. Adverbial clauses⁵ are also in most of the cases (92%) delimited by PP breaks, something which agrees with the theory's stipulations and the complement-adjunct distinction that is made (adverbial clauses being adjuncts). The clauses, on the other hand, which to a greater extent -with respect to other clause types- tend not to trigger a PP break at their left edge, are sentential complements. Their behaviour with respect to the other types (they compose 67 of total 73 instances in which a PBI 2 or 1 occurred) is fairly reasonable, given that they are more closely connected to the verb of their superordinate clause. Furthermore, in many cases the insertion of a phrase break at their left edge would result in a very short Phonological Phrase consisting only of the previous (i.e. the superordinate's clause) verb, as it would be the case for *says* in the next example.

(Rick Kahn)(says top city hall)(and police officials)(are engaged)(in a deadly game of denial)

Such short phonological phrases tend to be avoided. We will be discussing the relevance of length for prosodic phrasing in the next section.

3.2.1.2 SV Constituents

The vast majority of boundaries which were not predicted correctly, consists of subject-verb sequences, in which the subject does not form a Phonological Phrase on its own, but groups with the following verb instead⁶. This is actually a typical case of non-traditional constituency which has been used as an argument against the syntax-phonology connection due to the apparent anisomorphism. It is also an argument in favour of a flexible syntactic structure, such as Combinatory Categorical Grammar's, which allows for such non-traditional constituents. Whilst in CCG the grouping of the verb depends on factors related to discourse, there have been proposals (Gee &

⁵e.g. (There is no new library) (in sight),(*unless Dinhelman's money*)(*can be used*)(*for construction*

⁶According to AlignXP-R the right edge of the subject-maximal projection should align with the right edge of a phonological phrase. Furthermore, there is no Wrap-XP violation, since the subject does not surface within the VP. Finally, it should be made clear that these incorrect SV predictions involve lexical maximal projections. Otherwise (in the case of functional projections such as pronouns), the SV sequence is what the theory would have correctly predicted.

Grosjean 1983) accounting for such patterns on the basis of phrase length(reflected in the number of branches) aiming to achieve structural balance. In particular Gee & Grosjean (1983) predict the combination of the verb with its subject as opposed to its complement, when the former is a simpler prosodic unit than the latter. The complexity of its boundary is a function of 'the number of branching nodes dominated by the node dominating the boundary'. The flexibility of the verb in its adjoining direction is further theoretically supported on psycholinguistic evidence suggesting a certain freedom in the verb's patterning.

Even though Gee & Grosjean (1983) use this condition in an algorithm for predicting pausing performance, in order to group Phonological Phrases into Intonational Phrases, and further leave the possibility for semantics or discourse influence open, length is still in general considered to play a fairly important role in phonological phrase partition. In our corpus, in 187 SV patterns, 121 subjects were single words, 54 were two-word phrases, nine consisted of three words and there also were one four-word and six-word phrase (see figure 3.5).

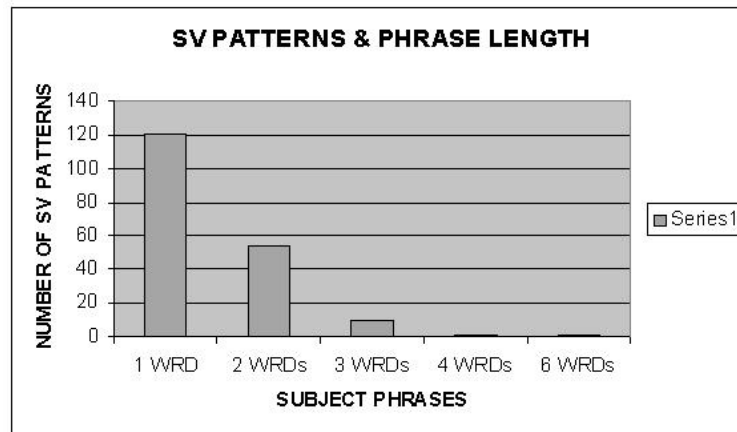


Figure 3.5: Occurrences of SV patterns with respect to the length of the subject

In all instances the head of the subject noun phrase coincided with the end of the noun phrase itself (no complement or adjunct structures following, which would cause the subject to be a heavier constituent in terms of length at least). Therefore, according to

these data at least, very short and structurally simple Phonological Phrases are generally avoided.

One thing that should be pointed out is that the majority of cases involve proper names. It seems that proper names display a behaviour more or less similar to that of pronouns in the sense that instead of forming Intermediate Phrases on their own, being a lexical projection themselves, they adjoin to adjacent maximal projections each time.

Within the discourse model, proper names are first introduced either explained (most of the times by an apposition structure) or constituting apposition structures themselves (for the familiarity of the hearer with the referent to be established), and are afterwards referred to unmodified. For example

Public Health Deputy Commissioner, *Van Dunn*, admits recent layoffs...*Dunn* says...

In the first case the apposition structure necessitates the formation of an independent Phonological Phrase⁷, while in the vast majority of the latter cases the proper name is not separated with a phrase break. One way to account for the latter is with respect to discourse. Since proper names in this case have already been introduced and are *old* discourse elements, they might lack the prominence, the information weight that would be signalled by phrase finality. On the other hand, the behaviour of proper names could be accounted for in terms of length again, since the insertion of PP break on their right edge would result in the formation of a one word phonological phrase.

3.2.1.3 Verb-adjunct sequences

In accordance with the domination and containment distinction⁸ a Phonological Phrase boundary was predicted between an adjunct structure and the previous maximal projection it attached to⁹. However, these kinds of phrase breaks did not always occur in

⁷Apposition structures along with non-restrictive relatives, parentheticals and pre-posed adverbials correspond to separate Intonational Phrases

⁸Constraints apply only on constituents dominated by an XP category, i.e. by all XP nodes(see section?)

⁹Adjunct structures are contained but not dominated by the XP.

SUBJECT				OBJECT		APPOSITION	
BREAK		NO BREAK		BREAK	NO BREAK	BREAK	NO BREAK
NMB	MB	NMB	MB				
22	23	2	112	-	7	37	6

Table 3.6: This table illustrates the presence of phrase breaks with respect to proper names occurring in a Subject position (that is whether a PB occurs at the right edge of the proper name), in an Object position (whether a PB occurs at the left edge of the proper name in accordance with the affiliation direction of object-pronouns) or as an apposition structure (in this case a phrase break always occurred at the right edge, thus only left edges are taken into account). NMB and MB stand for *not mentioned before* and *mentioned before* respectively

the actual data, especially in cases where the adjunct was immediately following the head of the projection that contained it. Compare for example the predicted and the actual partition for the following utterance

...(who is [[choosing words]_{VP} (*carefully*)_{VP} (predicted partition)

...(who is choosing words carefully) (actual partition)

We will be returning to that very point in the following section. Table 3.7 shows the distribution of phrase breaks before adjunct structures within noun and verb phrases in the corpus.

VERB PHRASES		NOUN PHRASES	
PHRASE BREAK	NO PHRASE BREAK	PHRASE BREAK	NO PHRASE BREAK
231 (77.77%)	66 (22.22%)	125 (83.89%)	24 (16.10%)
TOTAL (VPs AND NPs INCLUDED)			
PHRASE BREAK		NO PHRASE BREAK	
356 (79.82 %)		90 (20.17%)	

Table 3.7: Distribution of phrase breaks before adjunct structures. In 12 out of these 90 non-occurring phrase break instances, a phrase break was not predicted anyway because this prediction would result in the violation of Wrap-XP with respect to another constituent (e.g. cases of adjuncts in noun phrases that functioned as the verb's complement). Though it might seem that the claims of the theory with respect to Wrap-XP and its non-violation in the case of adjuncts are statistically relevant, it will be shown that there are other factors influencing the triggering of phrase breaks than a complement-adjunct distinction.

3.2.2 Phrase breaks not predicted

Table 3.8 illustrates the different cases of phrase breaks which occurred in the data but had not been predicted by the theory. Following are the examples for each type of false prediction starting from the top of the table. Underlined boundaries correspond to the ones not predicted.

(He dismisses)(claims)(by the Phoenix police)

(She will refer her former customers)(to the clinic)¹⁰

...(causing increased blood pressure)(tremors)(acne)(and other physical damage)

(The psychological)(effects of ice)

The category *others* includes mainly phrase breaks after functional words. The number of these cases, though, is insignificant considering the vast number of functional words in the corpus.

HEAD BREAKS	377	58%
DITRANSITIVE V. COMPL.	53	8.15%
COORDINATION STRUCTURES	64	9.84%
PRE-NOMINAL MODIFIERS	87	13.38%
OTHERS	69	10.61%
TOTAL	650	

Table 3.8: Phrase breaks not predicted. The category *DITRANSITIVE V COMPLEMENTS* includes 5 cases of object complements as well, e.g. (makes democracy)(more participatory).

3.2.2.1 PP breaks after phrase heads

One of the most typical cases in which the theory failed to predict existing phrase breaks are phrase breaks occurring after the head of a maximal projection (377 non

¹⁰The actual partition is in fact the following:
(She will refer)(her former customers)(to the clinic)

where there is also a non-predicted phrase break after the head *refer*, which corresponds to the first type in the table.

predicted PP breaks). Among the set of constraints on the formation of Phonological Phrases there is nothing which makes direct reference to heads or predicts phrase breaks with respect to them. Thus, the only way for the theory to account for such a case is for the edge of the head constituent to coincide with the edge of the Maximal Projection or for the phrase constituent following the head to be an adjunct structure. For example the predicted prosodic partition for the sentence *More petitions were certified in Massachusetts* would be the following:

(More petitions)(were certified)(in Massachusetts)

in which the NP head's edge aligns with the NP's edge and the VP head is followed by an adjunct. In all other cases the Phonological Phrase boundaries that occurred after phrase heads were not predicted. Tables 3.9 and 3.11 show the occurrence or non-occurrence of PP breaks within NPs and VPs and on the left edge of complements or adjuncts. In the former case, that of complements, no phrase break was predicted, while in the latter, that of adjuncts, it was. Subject complements (i.e. complements in copular constructions) are shown separately. Furthermore, the cases where the complement or adjunct structure was a whole clause were not taken into account, since the theory has been adjusted, so that for such cases a phrase break is predicted anyhow. Finally, since -as it has already been mentioned- the complement-adjunct distinction is not at all straightforward, tables are also provided for cases where the distinctions were easier to draw.

IMMEDIATELY FOLLOWING						NOT IMMED. FOLL			
COMPLEMENTS		ADJUNCTS		COPULAS		COMPLEMENT		ADJUNCTS	
BREAK	NO BREAK	B	NB	B	NB	B	NB	B	NB
168	305	57	44	12	140	48	10	174	22
35.5%	64.4%	56.4%	43.6%	7.8%	92.1%	82.8%	17.2%	88.8%	11.2%

Table 3.9: Complement-adjunct distinction in verb phrases

IMMEDIATELY FOLLOWING				NOT IMMED. FOLL			
COMPLEMENTS		ADJUNCTS		COMPLEMENT		ADJUNCTS	
BREAK	NO BREAK	B	NB	B	NB	B	NB
132	243	42	10	30	3	103	11
35.2%	64.8%	80.8%	19.2%	90.9%	9.1%	90.4%	9.6%

Table 3.10: Complement-adjunct distinction in verb phrases(Cases in which the distinction was easier to draw)

WRAP-XP VIOLATION				NON WRAP-XP VIOLATION			
COMPLEMENTS		ADJUNCTS		COMPLEMENT		ADJUNCTS	
BREAK	NO BREAK	B	NB	B	NB	B	NB
53	35	59	12	34	25	66	12
60.2%	39.8%	59	12	57.6%	42.4%	66	12

Table 3.11: Complement-adjunct distinction in Noun Phrases. For the cases under the Wrap-XP violation category, a phrase break was not predicted irrespectively of a complement-adjunct distinction, as such a prediction would cause a Wrap-XP violation with respect to another constituent (e.g. if the NP in question were a verb complement.)

WRAP-XP VIOLATION				NON WRAP-XP VIOLATION			
COMPLEMENTS		ADJUNCTS		COMPLEMENT		ADJUNCTS	
BREAK	NO BREAK	B	NB	B	NB	B	NB
39	24	49	7	29	18	59	8
61.9%	38.1%	87.5%	12.5%	61.7%	38.3%	88.1%	11.9%

Table 3.12: Complement-adjunct distinction in Noun Phrases(Cases in which the distinction was easier to draw)

First of all, an approximately forty and sixty percent for the occurrence and non-occurrence respectively of a PP boundary before a verb complement is not at all indicative of any correlation between the complement-adjunct distinction and the trig-

gering of phrase breaks, and consequently of any such effect syntax might have on Prosody. Furthermore, in noun phrases the situation is actually reversed and in most cases (58.38%) Wrap-XP is violated, as a phrase break is inserted after the NP head and before the noun complement with which it forms a single maximal projection. What is more, if one takes into consideration the 'immediately following' parameter along with the fact that complements are usually shorter than the adjuncts in the corresponding position, the complement-adjunct distinction seems even less important(cf. table 3.13).

IMMEDIATELY FOLLOWING		NOT IMMED. FOLLOWING	
BREAK	NO BREAK	B	NB
225(39.19%)	349(60.8%)	58(22.83%)	196(77.16%)

Table 3.13: The insertion of *phrase head breaks* as a function of the *immediately following* parameter.

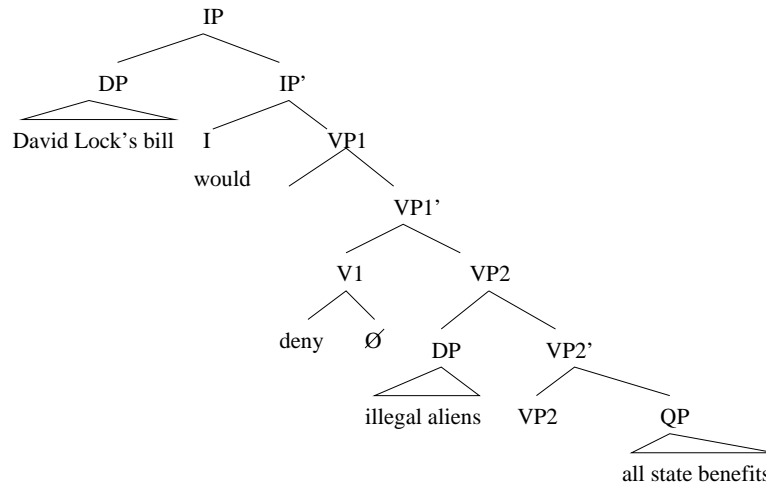
Adjuncts are more commonly separated with a phrase break because most of the times they do not occur in a *immediately-following the head* position (see table 3.14).

COMPLEMENTS				ADJUNCTS			
IMMED. FOLLOWING		NOT IMMED. FOLL.		IMMED. FOLLOWING		NOT IMMED. FOLL.	
620		58		248		196	
B	NB	B	NB	B	NB	B	NB
255	365	48	10	182	66	174	22

Table 3.14: Number of complements and adjuncts occurring in *immediately following* the head and *not immediately following* position.

Supporting the latter assumption is the case of ditransitive verbs. If Wrap-XP was valid, then both complements following the verb should be contained within the same as the verb phonological phrase, since both of them are dominated by the verb's maximal projection. However, the second verb, which is not immediately following the head, is in approximately 83 % of the cases separated with another phrase break (Fig-

ure 3.6). In this particular case ranking Align-XP/R higher than Wrap-XP would actually give the correct predictions.



(David Lock's bill) (would permanently deny)(illegal aliens)(all state benefits)

Figure 3.6: In 83% of the cases of ditransitive verbs, a PP break is inserted between the first (illegal aliens) and the second complement (all state benefits) resulting in a Wrap-XP violation (of VP1)

What seems to be the case here is that within every clause, each major constituent forms a separate phonological phrase (more like the PP partition following the increase in the event structure) with the constituent following the head of the phrase sometimes attaching to the former PP instead. Compare the following examples:

- (At present rates)(almost one hundred and fifty million children)(will die)(during the nineteen nineties)
- (Such publicity)(encourages)(gang members)(by feeding their hunger for status)
- (The media)(overstate)(the problem)(and most ignore his plea)(not to publicise)(gangs' names)

In any case, the fact remains that there is a significant number of 'head' breaks unaccounted for. One way to account for the non-predicted phrase breaks is with respect to speech style itself. It has already been mentioned that differences in speech style and rate may play a role in triggering phonological phrase breaks. In a not as fast plus a more clearly and finely prosodified, aiming at facilitating understanding, broadcasting speech, phrase breaks might appear stronger than in normal speech. According to Selkirk (1993), in English

the right edge of an independent(non-lexically governed) phrase is aligned with the right edge of a Minor Phrase(X_{ind} , R; MiP, R)

Along these lines the PP breaks found after lexical heads might be minor phrase breaks surfacing as PBs 3 or 4 in the corpus.

The occurrence, however, of prosodic boundaries is by no means uniform throughout the corpus. In not all cases do phrase breaks appear that strong at the edge of a minor phrase, while in many cases they appear too weak. To make this point clearer, table 3.15 shows the distribution of phrase breaks after verbs, where the edge of the phrase head in question does not align with the edge of the maximal projection and is not followed by an adjunct structure either. This way the MiP boundary will not according to the theory's stipulations coincide with a Major Phrase break, so the triggering of any phrase break will not be a result of this alignment. Furthermore, it should be made clear that most but not all of the VPs presented conform with the lexical government condition(The rest are included in accordance with the clause internal approach that was taken for the evaluation. For a more thorough investigation of Minor Phrases, the Accent Constraint should also be considered¹¹)

The majority of phrase breaks is almost evenly distributed among phrase breaks 3, 2 and 1. The explanation for the variation in the strength of the breaks may lie on the interaction of other factors such as the length of the phrase -either the current or the following one- assuming that the derivation of small phrases in this manner is discouraged.

¹¹A Minor phrase contains at most one pitch accent (Pierrehumbert & Beckman 1988)

	PHRASE BOUNDARY STRENGTH				
	PB INDEX 4	PBI 3	PBI 2	PBI 1	PBI 0
NUMBER OF PBs	67	114	143	134	3

Table 3.15: Minor Phrase Boundaries

There is, however, one proposal on the mapping of Phonological Phrases to Syntactic Structure which makes direct reference to the notion of a phrase head allowing also for the effect of length on the formation of prosodic units. Nespor & Vogel (1986) have proposed the following conditions on Phonological Phrases formation :

- A Phonological Phrase consists of a lexical head X and all constituents on the non-recursive side of the head's(X) phrase, which are contained in the same maximal projection. In English the non-recursive side is the left¹². The following example illustrates the partition of the utterance into Phonological Phrases according to this definition. The end of the Phonological Phrase is aligned with the right edge of a lexical head(*Chism, studied, poisoning, years*) and the Phonological Phrase is then extended to include all constituents on the left side of the head that are within the same maximal projection
(Julian Chism)(has studied)(lead poisoning)(for some 35 years)
- The Phonological Phrase may optionally be extended to include the constituent to the right of the phrase head as long as they are both part of the same projection. Across languages, this option may also be obligatory or forbidden. In general, this option may be determined by the speech rate or the length of the following

¹²Selkirk (1978) also made that proposal for English. In particular

The Phonological Phrase: Constituency

- An item which is the specifier of a syntactic phrase joins with the head of the phrase
- An item belonging to a non-lexical category joins with its sister constituent

e.g. (The absent-minded professor)(has been avidly reading)(about the latest biography)(of Marcel Proust)

the head constituent. For example in the sentence *The initial process worries Beard, who says it can offer simplistic solutions for complex problems* (see figure 3.4 for its corresponding parse tree) the initial partition would be the following (The initial process) (worries) (Beard), (who says) (it can offer) (simplistic solutions) (for complex problems)

but since *Beard* is a short Phonological Phrase within the same as the previous PP maximal projection (the VP), it may be attached to the latter and both form a single PP as in the next partition

(The initial process) (worries Beard), (who says) (it can offer) (simplistic solutions) (for complex problems)

In Italian, on which the proposal was primarily based, only 'non-branching' constituents were included in the previous Phonological Phrase.

The branching against non-branching distinction can be ultimately seen as a matter of length, since in the latter case the constituent is bound to be smaller. In short, what Nespor & Vogel (1986) call the restructuring of the Phonological Phrase, that is the inclusion of the following constituent, may be less frequent or banned for large constituents, and in addition less frequent for the broadcasting speech style in question. That implies that -in this test corpus at least- one should expect the appearance of many major phrase breaks after lexical heads.

Table 5.4 shows the evaluation of Nespor & Vogel's (1986) proposal for our corpus.

PREDICTED CORRECTLY	NOT PRED.	PREDICTED INCORR.	RESTRUCTURING
2547	126	312	621

Table 3.16: Evaluation of Nespor and Vogel's proposal

The fourth column refers to instances in which no phrase break occurred at the edge of the head constituent, but they could nevertheless qualify as cases of restructuring and cannot therefore be considered as an indication of the proposal's failure.

Taking a closer look at the cases of restructuring, the following should be pointed out: First of all, length appears to play a significant role here as well, in accordance with the theory's stipulations. Non-branching structures following the head, that is to say single words, underwent restructuring in all (162 in total) but four cases. Moreover, function and content word sequences, which are reasonably the next shorter constituent, were categorised as follows: In 189 cases restructuring occurred, whilst in 62 cases it didn't. Finally, in order to assess how longer phrases behave with respect to restructuring, three word and four or more words sequences were examined. Table 3.17 shows the results and table 3.18 and graphs 3.7, 3.8 summarise all the above data.

	RESTRUCTURING	NO RESTRUCTURING
3 WORD SEQ.	155	161
	49%	51%
4 OR MORE WORD SEQ.	73	155
	32%	68%

Table 3.17: Restructuring in three or more words sequences

	RESTRUCTURING	NO RESTRUCTURING
1 WORD	158	4
	97.5%	2.5%
Funct. WRD and Cont. WRD	189	62
	75.3%	24.7%
3 WORD SEQ.	155	161
	49%	51%
4 OR MORE WORD SEQ.	73	155
	32%	68%

Table 3.18: Restructuring as a function of the following phrase's length

It might worth mentioning that a great number of cases in which restructuring occurred involved phrasal heads which inherently lack informational content, ranging from link

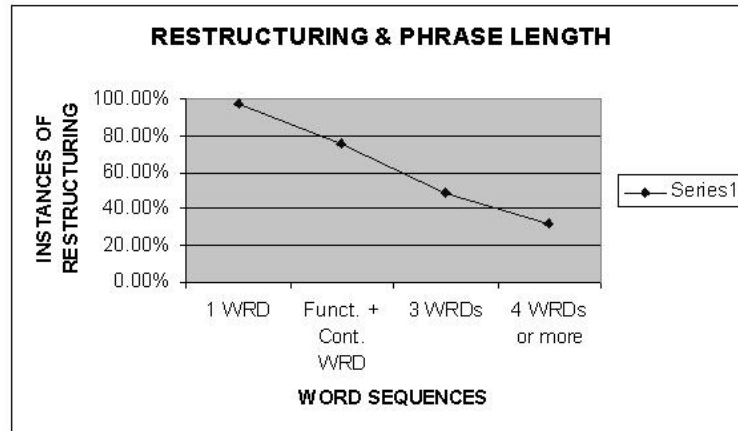


Figure 3.7: Percentage of restructuring cases for phrases of different length

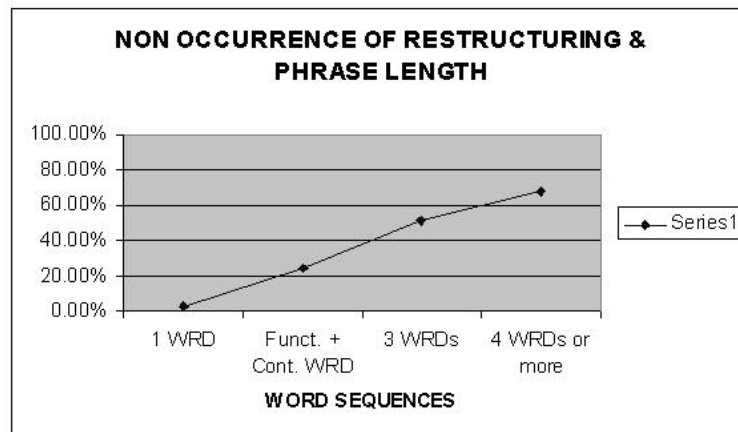


Figure 3.8: Percentage of cases in which restructuring did not occur for phrases of different length

verbs (mainly the verb *to be* and further verbs such as *become*, *considered*, *feel* etc.(cf table 3.9, copulas)) to semantically light verbs such as *make*, *do*, *get*. All these verbs -except for few instances- displayed the same behaviour with respect to restructuring. Similarly, idioms and expressions in which the head of the phrase is deprived of 'individual' meaning (and is more closely related to the rest of the phrase) underwent restructuring as well (There were, however, very few instances of these in the data). In general, semantic weight and informational content have been proposed to affect

the choice of focus (see for example Ladd (1996)) and considering that accentuation is usually associated with phrase finality and ultimately may be regarded as a function of prosodic phrasing, lexical information about individual words could be used to improve the proposal's performance. The variable of semantic weight and informational content could further be used in order to achieve a unified account of functional and content words with functional elements, copulas and semantically light words on the one edge, semantically heavy and focused constituents on the other.

Finally, there is one last point to be made with respect to the empirical data Nespor & Vogel's (1986) observations are based on. Their corpus corresponded to a normal rate, colloquial speech style, which is fairly different from that of the BURadio Corpus. Apart from what implications that may have for the effect of speech rate and style per se, it forms an argument against the first attempt to account for the occurrence of 'head' breaks as Minor Phrase breaks surfacing stronger in a different speech style, since this occurrence has already been observed in a normal rate, colloquial speech style. Note, however, that Nespor & Vogel's (1986) data are still read speech (*speakers reading and/or imitating, in as natural wa as possible* (Nespor & Vogel 1986)).

3.2.2.2 Coordination structures

Another observation that Nespor & Vogel (1986) made and is relevant to our discussion is the one concerning lists and 'list-like' structures. Such structures form the second type of phrase breaks that were not predicted. By lists we refer to sequences of constituents of the same type, in our case in particular asyndetic structures and conjunctions. According to Nespor & Vogel (1986)

In a sequence of more than two constituents of the same type, i.e. X1, X2...Xn, an intonation break may be inserted before each repetition of the node X, i.e. before X2, X3 ...Xn.

The following underlined coordination structure is an example from the test corpus
 ...a tax (with Boston's share) (earmarked) (for public safety),(youth outreach programs)
(and educational reform).

Assuming that coordinated constituents are dominated by a functional conjunction phrase (see figure 3.9) rather than projecting to phrases of the same type, the triggering of boundaries should not violate Wrap-XP, unless the whole conjunction phrase is dominated by another lexical maximal projection (as for example would be the case if the conjunction structure were the complement of a verb).

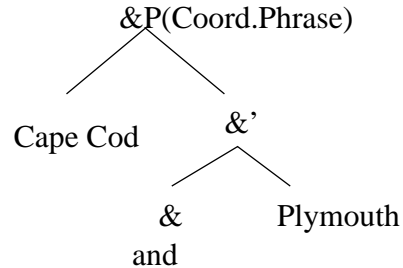


Figure 3.9: Coordination Structures

Coordination structures, however, behave similarly irrelevant of any Wrap-XP violation or non violation, and it is the former case in which non-predicted by the theory phrase breaks occur. Table 3.19 shows the distribution of phrase breaks for each type of constituent and conjunct. In approximately 85% of the total number of coordination structures in the corpus there is a phonological phrase break inserted according to Nespor & Vogel's (1986) observations (73% corresponding to a Full Intonational Phrase boundary as well). One thing that should be pointed out is that longer coordinated phrases such as verb phrases or whole clauses display a consistency with respect to the triggering of phrase breaks. In fact, in all cases they are produced with their own Intonational Phrase. On the other hand shorter constituents like noun phrases (in the cases at least where there was a weaker boundary(PB2/1)) and particularly adjectives are less consistent as far as PP triggering is concerned.

PHRASE TYPE	CONJUNCT TYPE	PBI 4	PBI 3	PBI 2	PBI 1/0
NP	and	26	21	10	16
	but	2	-	-	-
	or	6	3	-	-
	-	28	4	-	-
SUBTOTAL		70(56.5%)	28(22.6%)	10(8.06%)	16(12.9%)
Adj.P	and	2	3	1	5
	but	-	-	-	-
	or	1	2	-	1
	-	-	1	-	-
SUBTOTAL		3(18.8%)	6(37.5%)	1(6.2%)	6(37.5%)
VP	and	12	-	-	-
	but	2	-	-	-
	or	1	-	-	-
	-	2	-	-	-
SUBTOTAL		17	-	-	-
Clause	and	56	-	-	-
	but	32	-	-	-
	or	1	-	-	-
	-	2	-	-	-
SUBTOTAL		91	-	-	-
TOTAL		181(73%)	33(13.3%)	11(4.4%)	22(8.9%)

Table 3.19: Coordination structures

3.2.2.3 Pre-nominal modifiers

In Section 3.1 it was mentioned that for the purpose of evaluating Truckenbrodt's (1999) theory Wrap-XP was ranked higher than Align-XP/R. If Align-XP/R was ranked higher instead, the resulting prediction would suggest the existence of a phrase break after every single maximal projection, irrespectively of its function -whether it is for

example a pre-modifier or a complement following the phrase head. And while it can be concluded by the data that in the case of maximal projections following the head of a phrase (which they were or were not dominated by), the distribution of phrase breaks could accommodate a reversal in the ranking of the two constraints with no serious effect on the theory's predictions¹³, the case of elements that are traditionally characterised as modifiers preceding the head of a syntactic phrase has so far been 'conveniently avoided'. In earlier stages of X-Bar theory adjective phrases or noun phrases preceding a NP head were considered to be noun modifiers which were included within the noun's projection. The formulation of constraints such as Wrap-XP or the lexical government constraint would among other things block the prediction of phrase breaks after these pre-modifier phrases, something which would be in accordance with empirical observations and consequently would necessitate that these constraints are ranked higher than any alignment constraint.

In evaluating the theory, that traditional view of the noun phrase was adopted. However, the most recent proposals on DP and NP structure suggest a model in which traditional pre-modifiers -adjectives in particular- are specifiers of functional projections inserted within the DP but outside the NP and thus not affected by the Wrap-XP constraint (see figure 3.10).

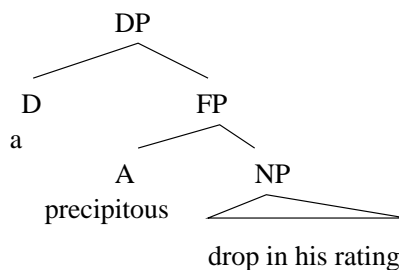


Figure 3.10: DP Structure

¹³Ranking Align-XP/R higher would for example correctly predict the 48 insertions of phrase breaks after the first complement in ditransitive constructions or the 59 instances of PBs before adjuncts in NPs and 64 PBs in coordination structures which were not predicted due to other X-Wrap violations, but it would still over-generate in the corresponding cases, where there was no phrase break and more importantly still not account for all the instances of phrase breaks after phrase heads.

The way Truckenbrodt's (1999) set of constraints is formulated, nothing would prohibit the insertion of phonological phrase breaks after the modifying phrase¹⁴ and one would end up with a very large number of falsely predicted phrase breaks. In all thirty news stories of the test corpus, there were only 83 instances of Phonological Phrase boundaries occurring after maximal projections in pre-nominal position. Just as an indication of the theory's potential performance -if the more traditional view of the noun phrase structure had not been taken- in story one alone of the corpus, there would have been 13 more phrase breaks incorrectly predicted if Wrap-XP was ranked higher and 27 phrase breaks incorrectly predicted if AlignXP-R was ranked higher. In general, ranking Align-XP/R higher would yield slightly better predictions (see footnote 13), if it weren't for maximal projections in pre-modifying position.

As far as the 83 instances of pre-nominal phrase breaks are concerned, it should first of all be mentioned that most of them occurred within complex DPs consisting of several maximal projections. In particular, in only nine cases out of the 83 is there a phrase break inserted within a simple one-word modifier plus noun sequence (e.g. (a welcome)(change)). This would suggest again an effect of length, while it could also be viewed as an effort to facilitate comprehension and processing of the noun phrase. For example, in the following DP, the phrase break after *causing* facilitates the division of the phrase into the correct syntactic-semantic units

(cancer causing)(asbestos fibres)

Similarly

(the Massachusetts)(Auto Body association)

(low quality)(dark colored glass) etc.

In many cases the insertion of a phrase break involves a distinction between attributive, non-restrictive adjectives and restrictive noun modifiers. For example

(same)(ideological cloth)

(the two term)(district attorney)

Compare the two following cases as well

¹⁴The predicted phrasing for the phrase in figure 3.10 would be (a precipitous)(drop in his rating)

(fetal protection policy)

(similar)(fetal protection policies)

In general, though, complex DP structures only set an option, a possibility for the triggering of a phonological boundary within them. Typically, however, there is no PB insertion, and pre modifiers-NP heads groupings are in fact the only construction which consistently allows for longer Phonological Phrases, e.g.

(obsolete public library)

(Merrill Lynch downtown Boston office)

Finally, it should be pointed out that Nespor & Vogel's (1986) theory runs into the same problems with respect to the way the DP is structured. Recall in page 42 that according to Nespor & Vogel (1986)

A Phonological Phrase consists of a lexical head X and all constituents on the non-recursive side of the head's(X) phrase, which are contained in the *same maximal projection*. In English the non-recursive side is the left.

Pre-nominal modifiers can thus no longer form a single Phonological Phrase within the head of the noun phrase, since they are no longer within that NP. It seems that this problem, which all the syntactically motivated proposals are facing, is just an instance of what may be the most significant mismatch between syntactic and prosodic structure. The syntactic structure is a much more complex one, with finer distinctions, more boundaries, nested constructions and deeper embedded constituents.

One way to overcome that particular problem, though, might be if the conditions on the domain of the PP were stated in terms of extended projections (Grimshaw 1990) of noun phrase heads at least, rather than in terms of maximal projections. A noun has an extended projection into a DP, as a verb has an extended projection into an IP. The VP's extended projection, however, would then also include the subject, something which we would want to avoid, given that subjects do form separate PPs in most cases. For example, in the figure 3.11 the DP *Similar policies* is the extended projection of the noun *policies* and the IP *Similar policies were struck down* is the extended projection of the verb *struck*. If Nespor & Vogel's (1986) definition of the Phonological Phrases

was changed to refer to noun phrases in terms of extended projections separately, i.e. for $X=N$, the phonological phrase triggered by X consists of all constituents on the non-recursive side of X , which are contained in the same *extended* projection, the adjective *similar* would be included in the same as the NP phonological phrase, yielding the following prosodic partition
 (Similar policies)(were struck down)
 as opposed to
 (Similar)(policies)(were struck down)

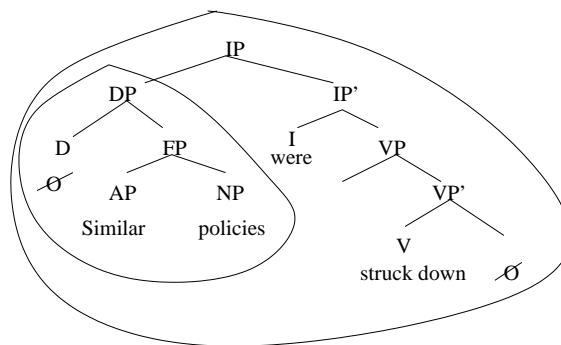


Figure 3.11: Extended projections

Note that the notion of extended projections would not prove as useful for Truckenbrodt's (1999) set of constraints, since the complement-adjunct distinction, which allowed it to predict more finer partitions found in the data, could no longer be maintained¹⁵.

Finally, note that such extended projections are, however, conspicuously reminiscent of types used in the semantically based formalism which we will be turning to in the very next chapter, i.e. Combinatory Categorical Grammar.

¹⁵Since, in order to be of use in our case, it can only be stated in terms of lexical maximal projections, the situation would be further complicated.

Chapter 4

Evaluation of M. Steedman's proposal

4.1 Information Structure Identification in complex texts

The interaction between Intonation and Information Structure can be approached under two different scopes within the CCG formalism. In the first case, prosodic information is used to resolve and justify the so-called spurious syntactic ambiguity of CCG resulting from a proliferation of alternative derivations for a single sentence; and in the second case, which interests us here, Syntax -Information Structure to be exact- is inversely used to predict Intonation. Therefore, the problem of predicting Prosody turns out to be an Information Structure Identification task. Unfortunately, the problem is now hardly reduced. Identifying Information Structure is a very complicated issue in its own right, especially if one considers the different theories and schools of thoughts on the properties and domain of IS and even more importantly the fact that it involves notions such as the speaker's and the hearer's attentional state and prior knowledge, presupposition, accessibility, salience, conversational implicature and so on. Only the different terminologies¹ that have been used, in addition to the fact that there does not seem to be a widely accepted definition of IS are indicative of the state of diversity within IS theory. The task of IS Identification becomes even more difficult,

¹e.g. theme-rheme, given-new, contextually bound and non bound, topic/ground/presupposition-focus.

if one considers the complexity of the utterances in our test corpus². There have actually been comparatively few attempts to apply an IS partition on realistic texts rather than 'tailored' examples (Hajicova, Skoumalova & Sgall (1995), Hahn (1995), Hoffman (1996), Komagata (1998)).

Since there really are not any clear-cut definitions for the properties of IS parts and no widely accepted guidelines on how to identify them, plus tests -such as the question test³-to uniquely identify an IS partition, are hardly applicable in real texts, evaluation is bound to be difficult and subjective. In order to make the whole evaluation process more reliable only particular parts within the news stories, where the identification of IS was clearer, were chosen.

4.2 Criteria for Identifying Information Structure

First of all, it should be mentioned that our approach to identify Information Structure, in order to evaluate Steedman's (2000*b*) proposal, reasonably is based on his two-dimensional view of Information Structure (see Section 1.5) with particular emphasis on the first dimension, a theme-rheme partition underlined by an informational contrast between the two IS components.

In accordance with that view the specific criteria for recognising Information Structure would be (following Komagata (1998)) the following :

²This is how the first story in the corpus begins. It is also a typical set of sentences throughout:
A nineteen eighteen state constitutional amendment made Massachusetts one of twenty three states where citizens can enact laws by plebiscite. The first ballot initiative adopted by Bay Staters was a nineteen twenty measure that defined cider and beer as non-intoxicating, exempting them from prohibition. The task gets more complicated, when reported speech is also involved
...Heightened awareness of the process may be one reason why more petitions were certified this year in Massachusetts than ever before, says the Secretary of State's Chief Legal Counsel, David Sullivan.

³Information Structure can be identified in a question-answer context or in our case by finding an appropriate question, the response to which is the utterance

Referential Status Information Structure has often been related to notions such as contextual link, contextually bound and non-bound, given vs new, all of which refer to the discourse status of IS components, whether these are entities or abstract objects(events, states etc.) associated with noun phrases or verb phrases, SV sequences and clauses respectively. Theme is primarily characterised as contextually linked to the previous context, as opposed to rheme, which is a new element advancing the discourse. The rheme, however, can be a contextually bound element as well, and advance the discourse similarly. Thus whilst IS cannot be reduced to referential status, referential status can be used as a good indication for the theme-rheme partition. We can further specify what exactly we refer to as contextually linked by making use of Prince's (1981) given-new information taxonomy. A contextually linked element should fall then under the evoked and inferable categories⁴.

Contrast Contrastiveness is defined as a relation between different referents made available from the context(discourse and situational context). Contrast is again a property which characterises both theme and rheme, even though it is more explicit for the rheme part of the utterance. That is while rheme has to be contrastive, the theme does not.

Linguistic form Surface structure, word order and lexical meaning have been shown to affect and mark Information Structure. Theme is usually found in utterance-initial position in accordance with the old-information first principle; Hajicova et al. (1995) use the verb's lexical meaning at one step of their algorithm to identify it as topic(general lexical meaning) or focus(specific lexical meaning); and certain linguistic structures such as cleft and pseudo-cleft structures, extraposition, left dislocation, inverse word order⁵, may also be considered to mark IS

⁴An entity which is already in the discourse model is evoked. A discourse entity which can be inferred via logical or plausible reasoning is inferable

⁵Examples of such structures (from Quirk, Greenbaum, Leech & Svartnik (1985))

It is her vegetables that Julie buys at the supermarket (cleft)

What you need most, is a good rest (pseudo-cleft)

To settle the matter was my prime objective (extraposition)

Your friend John, I saw *him* last night (left dislocation)

Faint grew the sound of the bell (inverse word order)

partition.

4.3 The choice of Information Structure domain

Another issue that should be taken into consideration is the domain of Information Structure, the structural level at which Information Structure distinctions are made. Even though it is generally thought to be utterance-internal there are diverse opinions regarding the organisation of information within the utterance level (and thus the domain itself). The three different approaches can be summarised to:

- Traditional view: Information Structure is a non recursive matrix level organisation. There is a single *binomial* theme-rheme distinction accounting for the whole utterance. Steedman's background -focus partition can be accommodated within this framework, to account for finer distinctions in information organisation. In another approach taken by Chafe, Vallduvi, Burring, Hoffman and others, Information Structure is defined in terms of a *trinomial* partition in Link-Focus-Tail, which substitutes for the notion of discontinuous themes in Steedman's (2000*b*) approach (for an explanation and comments on this notion see next section).
- Information structure may consist of multiple theme-rheme partitions, that is partitions within the main and subordinate clause in a complex utterance, resulting in the clause being the domain of IS. For example, Kruijff-Korbayova & Webber (2001), in the interpretation of *although* as a discourse connective, assume the following IS organisation:
(Although Clyde married)_θ(BERTHA)_ρ(he)_θ(did not inherit a PENNY)_ρ
- The organisation of Information Structure can be recursive according to some researchers. Partee (1996) proposes the possibility of recursion in IS at least at the level of full embedded clauses. The domain of IS would thus be both the utterance and the clause.

- Communicative dynamism (Firbas 1964). This approach clearly separates itself from the rest, as the Information Structure is no longer characterised in terms of a simple binary distinction into theme and rheme. Instead, each element of the utterance is placed on a hierarchical scale according to the extent, degree of its contribution to the advancing of the discourse.

The approach taken in Steedman's theory seems to be in line with the traditional one. His two-dimensional view of Information Structure (see Chapter 1), though, seems to be compatible with the second (clause internal IS organisation) as well. As far as recursivity is concerned, focus is not suggested to be recursively projected. Finally, a clear-cut, binary theme-rheme partition cannot be accommodated within a scalar communicative approach, being a categorical distinction. Out of the first two proposals on the domain of IS, the second (clause-internal) empirically at least seems to be a better fit for our data, since it allows for a more detailed division of the utterance into information units. In the following utterance, for example, an IS partition within the subordinate clause can account for the phonological boundary found after the subordinate's subject.

(But when the fetal protection policy took effect)_{θ1}, (Green)_{θ2}(was transferred to a low lead area)_ρ (Matrix level IS partition)

(But when the fetal protection policy)_θ(took effect)_ρ, (Green)_θ(was transferred to a low lead area)_ρ (Clause internal IS partition)

(But when the fetal protection policy)(took effect),(Green)(was transferred to a low lead area) (Phonological Phrase partition).

Therefore, the second approach will be followed , whenever such a division seems possible and compatible⁶.

⁶It should be noted that the same 'detailed' division could presumably be accounted for, using disjunctive themes and rhemes which are directly allowed in Steedman's (2000*b*) proposal

4.4 Evaluation Results

The basic problem that the theory faces seems to be that exact incompatibility between a fairly simple, usually binary IS partition and a finer multiple Phonological Phrase partition. For example, consider the IS and PP division for the following utterance (in a context of *Marshall saying things*)

(Marshall says)_s (it would take about ten million dollars to equip the building for glass making with six furnaces operating round the clock to melt down truckloads of asbestos waste each day)_p

(Marshall says)(it would take about ten million dollars) (to equip the building)(for glass making)(with six furnaces)(operating round the clock)(to melt down)(truckloads of asbestos waste)(each day) (PP partition)

As an indication of the mismatch the following table shows the results of the theory's evaluation for 167 sentences in the test corpus. The evaluation was based on a clause internal approach to the organisation of IS, whenever such finer distinctions were supported by the context.

PP BREAKS PRED. CORRECTLY	PP BREAKS PRED. INCOR.	PP BREAKS NOT PRED.
489	89	323

Table 4.1: Evaluation of Steedman's theory

First of all, it is quite obvious that there is a significant number of non-predicted phrase breaks, which basically verifies the point made above about the inherent contrast between a simple IS organisation and a complex PP partition.

Furthermore, the instances of phrase breaks predicted incorrectly are not to be taken as a failure of the theory. That is because, according to Steedman's (2000*b*) claim at least, the grammar should allow the postulation of phonological boundaries, even when they are not found in the actual utterance. Such boundaries can be licensed by instances of unmarked themes or by cases in which unification between words with different information feature values is not possible.

Unmarked themes are themes which are not contrastive. According to Steedman (2000*b*) -following Pierrehumbert's notation-

If the rheme alternative set that a theme presupposes is unambiguously established in the context, it is common to find that theme de-accented throughout...without any pitch accent or obvious boundary.

Due to the lack of an explicit phonetic boundary, utterances carrying unmarked themes are bound to be ambiguous with respect to Information Structure. For example the following utterance (from Steedman (2000*a*))

Marcel proved COMPLETENESS
H* LL%

could be a felicitous response to both the following questions, each one indicating a different IS partition:

Which result did Marcel prove?
(Marcel proved)_θ (completeness)_ρ.
What do you know about Marcel?
(Marcel)_θ (proved completeness)_ρ.

For our purposes, what these examples imply is that an IS division may not appear phonetically implemented. Things get even more complicated, if one takes into consideration the inherent ambiguity of Information Structure particularly in lack of clearly disambiguating context fostered by the possibility of theme accommodation as well. In Komagata (1998) all three following utterances are argued to be 'legal' responses 'as long as the theme is linked to the context and the complementary rheme is a projected contrast...Although the contextual force of a question is very strong, it cannot completely specify the response. There is room for the respondent to accommodate a distinct theme.'

Q: Who did Felix praise?

A1: (Felix praised)_θ (Donald)_ρ

A2: (Felix)_θ (praised Donald)_ρ

A3: (Felix praised Donald)_ρ

To make these points clearer, take for example all the instances of proper names combining with verbs that were found in our data. For instance, in the next utterance :

Green was fifty years old at the time

there was no edge tone after the proper name, as it was shown to be the case most of the other times as well (see Chapter 3). The utterance occurred in the following context.

(Green)_θ (was transferred to a low lead area)_ρ . (Her base salary)_θ (remained the same)_ρ
(but she)_θ (lost incentive pay)_ρ

(Green)_θ (was fifty years old at the time)_ρ

If the IS partition stands as it is marked, then a phonological boundary has to be hypothesised after the proper name. That is also the case with all the pronouns (such as *she* in the utterance) serving as the theme; a boundary has to be hypothesised every time. Consider further the following discourse:

Utt1: Flynn offered a remedy that he admits is not likely to be popular.

Utt2: Flynn said with local option taxes such as an increase in sales tax the money goes to the community that generates it.

Utt3: For example a one cent sales tax increase said Flynn would add thirty five million dollars a year to the city's treasury.

Utt4: The mayor said soon he'll seek legislative approval of such a tax with Boston's share earmarked for public safety youth outreach programs and educational reform

For the first utterance we have assumed a context about *Flynn doing things*. For the second utterance there could be two IS partitions

(Flynn)_θ(said with local option taxes such as an increase in sales tax the money goes to the community that generates it)_ρ.

(Flynn said)_θ(with local option taxes such as an increase in sales tax the money goes

to the community that generates it)_p

The first one would be in accordance with the theme-set in the previous utterance. The second one would be based on utterances that occurred previously in the discourse⁷, and would also be in accordance with what follows (a new context about *Flynn saying things*). In all utterances, there is no phonological phrase break after the proper name. If we adopt the second partition for Utt2, the phonological boundary, which occurs after *say* is correctly predicted; if we adopt the first partition, a phonological boundary after the proper name needs to be hypothesised. In each case the theory's predictions turn out to be literally unfalsifiable⁸.

Since the task of identifying IS structure is rather difficult, in order to facilitate it, and furthermore test additional claims that the theory makes about the discourse meaning of tones, the distribution of pitch accents was also briefly examined. For one, there does not seem to be a clear correspondence between accent and its θ - or ρ - function. The marking properties of pitch accents are:

L*+H, L+H*: theme

H*, L*, H*+L, H+L*: rheme.

In the following utterance⁹, for example *policy*, which is contextually linked and in sentence initial position is, nevertheless, ρ -marked, whilst e.g. *humiliating* and *infuriating*, which are comments to *policy* are θ -marked. Note, furthermore, that the first name in *Virginia Green* is θ -marked, while the last name is ρ -marked.

⁷Boston Mayor Ray Flynn vehemently denied charges that there is a double standard in the city, that white life is more important than black life. The problem, said Flynn, is that there aren't enough financial and human resources to protect the public safety and improve the quality of life

⁸Note Ladd's (1996) criticism on the inclusion of boundaries that are not phonetically present in Steedman's analysis

...it is clear that the phonetic accountability of his analysis is compromised by its dependence on entities whose presence cannot be independently verified.

⁹The previous context to these utterances is:

He was president of the United Auto Workers Union Local in nineteen eighty-two, when Johnson Controls instituted a new policy. It banned fertile women from working in areas where they would be exposed to lead, in order to protect fetuses, should the workers become pregnant.

The policy also is humiliating says Joanne Leard

H* L*+!H L+H* L_L% H* L+H* L_L%

a union health and safety monitor who's worked at Johnson Controls for sixteen years

And Virginia Green says in her case the policy is infuriating

L+H* !H* L_ H* L_L% H* L+H* L_L%

Secondly, the distribution of pitch accents causes again a lot of IP boundaries to be hypothesised. Pitch accents set the INFORMATION feature of the word they are associated with to *theme* or *rheme* according to each accent's discourse function. These words may then combine with other syntactic categories as long as the value of their INFORMATION feature is unifiable. Thus θ -marked constituents may only combine with other θ -marked constituents or constituents whose INFORMATION feature value is left unspecified (supposedly carrying a null tone (absence of any tone)). And the same thing of course applies for ρ -marked constituents. In the sentence *but the state's most famous initiative was proposition two and a half* (see next page) for example (apart from the apparent lack of correspondence between the proposed IS partition and the projection of pitch accents) a phonological boundary needs to be hypothesised after the word *famous*. The L+H* accent on *famous* sets its INFORMATION feature value to theme, in contrast to the H* and L* tones which mark the rest of the words they are associated with as rheme. Thus, *famous* cannot combine with *initiative* since their INFORMATION feature values are not unifiable. Accordingly, 'themehood' or 'rhemehood' cannot be further projected and a boundary needs to be hypothesised after *famous* to delimit the projection of theme and mark the theme part of the utterance. Unification on the contrary is possible for the last part of the utterance *was proposition two and a half* and the range of the rheme's projection is normally delimited by the existing boundaries.

Even though within the theory these discontinuous -and yet adjacent- themes or rhemes could combine into a single theme or rheme, there is still no clear explanation why such a partition would occur or how it could be predicted.

Finally, to give the reader a better idea of the problems and issues in question, a short news story is presented below. The first paragraph corresponds to the proposed Information Structure partition, while the second one illustrates the projection of pitch accents according to their proposed marking properties. In both paragraphs the phonological partition is denoted by square brackets.

([On one side of the debate][before the Health Care Committee])_θ([are those who claim][that resident physicians][need legal protection])_ρ. ([Residents]_θ(routinely work][ninety][to one hundred and ten][hours per week])_ρ. ([Cambridge Hospital Intern])([David Ray])([applauds])_ρ([the proposal][to cap the work week][at seventy two hours])_θ. ([On the other side of the controversy])_θ, ([are training directors][from several hospitals])_ρ([who testify][that they are voluntarily][improving][residents' working conditions],[and don't need][a legislative mandate])_ρ. ([Sponsor of the bill])_θ, ([Boston Representative John McDonnough]),([says he does not expect]_ρ (the measure])_θ([to win approval][this session])_ρ. ([But McDonnough says]_θ(the bill])_θ([puts hospitals on notice][to take better care][of their doctors][in training])_ρ. ([For WBUR])_θ([I'm]_θ(Margo Melnicove])_ρ.

[On one_ρ side of the debate]_ρ [before the Health_ρ Care Committee][are those_ρ who claim]_ρ [that resident_ρ physicians]_θ [need legal_θ protection]_θ. [Residents_ρ routinely work]_θ [ninety]_ρ [to one hundred_θ and ten_θ][hours per week]_θ. [Cambridge_ρ Hospital Intern]_ρ [David Ray]_ρ [applauds]_ρ [the proposal]_ρ [to cap_ρ the work_ρ week] [at seventy two_ρ hours].[On the other_ρ side_ρ of the controversy]_ρ, [are training_θ directors][from several_ρ hospitals]_ρ, [who testify]_ρ [that they_ρ are voluntarily]_θ[improving]_ρ[residents'_ρ working_ρ conditions], [and don't_ρ need]_ρ [a legislative_ρ mandate]_θ. [Sponsor_ρ of the bill], [Boston_θ Representative John_ρ McDonnough]_ρ, [says]_ρ [he_ρ does not_ρ expect_ρ the measure][to win_θ approval]_θ[this session]_ρ. [But McDonnough_ρ says_ρ the bill]_ρ [puts

hospitals, on notice] [to take better, Care] [of their doctors] [in training]. [For
WBUR] [I'm, Margo Melnicove].

Chapter 5

Conclusions, Applicability and Future Directions

To continue where we left off, it seems that the CCG approach is the most problematic, if nothing else, at least with respect to its evaluation process. The basic assumption, however, which underlies the whole approach, that is the dependence of Prosody on the information organisation of the utterance, cannot just be dismissed as irrelevant. For one, discourse is more or less generally accepted to affect the intonation contour and determine the acceptability or not of an utterance's prosody. The acceptability or not, for example, of the following response depends on the established context.

Q: What did Marcel prove?

A: (Marcel proved)(COMPLETENESS)

Q: Who proved completeness?

A: #(Marcel proved)(COMPLETENESS)

The issue, however, is not whether discourse is relevant or not, but the extent of its relevance in determining Prosody; that is whether it exclusively determines it, whether the correspondence between the two structures is that straightforward and if such is the case, whether the approach proposed by Steedman (2000*b*) is the correct way to try and define that correspondence.

As far as pitch accents and their proposed function as ρ - or θ - markers are concerned, in more than several instances that proposal was found not to stand. But the identification of pitch accents or even the inventory of available and distinguishable tones, is still a matter of dispute among phonologists and phoneticians. The interannotator agreement for the ToBI labelling has yet to reach a really satisfactory high percentage, making it hard to determine which of the theory's failures are due to an annotation problem or a problem with the theory's actual claims. The whole issue becomes more problematic, if one considers that the accents most easily confused in the ToBI labeller consistency study (see Chapter 2) were the H^* and $L+H^*$ accents, which bear different θ/ρ marking properties. It would, thus, be useful to assess the proposal, using a parametric representation of F_0 contours such as, for example, the Tilt model¹

As far as the phonological phrasing itself is concerned and its -compared to a simple binary IS partition- increased complexity, the proposal could benefit from a better understanding and a more accurately determined Information Structure. The Identification Problem is of course a disadvantage for the proposal's applicability at least, but not directly for the validity of the claim itself. Tackling the Identification Problem for complex sentences, might turn out to bridge the gap between Prosodic and Information structure inserting finer distinctions in the latter. 'Meanwhile' it might be worth testing the proposals on less complicated structures and particularly dialogic texts.

In general, however, the mismatch between the theory's predictions and the actual phrasing was significantly large. Apart from that, what might be the most important argument against this 'one-sided' proposal is that other factors such as length seem to have a considerable influence on prosodic structure.

The more syntactically based approaches such as Truckenbrodt's (1999) or Nespor & Vogel's (1986) are at least easier to implement and evaluate. What is more important, though, is that within Optimality Theory these proposals allow for the possibility to

¹In Taylor (2000), it was shown that there was a significant degree of overlapping among ToBI tone annotations represented in Tilt space

include all various types of conditions², syntactic, semantic or purely phonological, in order to account for the norm and the variability in the data. Truckenbrodt (1999) for example proposes another constraint which is related to discourse, namely ALIGN-FOC

ALIGN-FOC = ALIGN(Foc, R; P,R) ‘Each focused constituent is right-aligned with a p-boundary’

And in contrast to a purely Information Structure based explanation of prosodic phrasing, all these various types of constraints have been shown to apply. Defining what their exact interaction is could allow then for a broader coverage of the data.

For phrasing prediction with respect to this test corpus at least, length was shown to play an important role, both in Nespor & Vogel’s (1986) restructuring of the Phonological Phrase and the occurrence of non-traditional Subject-Verb constituents. Very short Phonological Phrases were shown to be avoided, as well as very large ones, in the sense that a very large syntactic projection would be broken down in more than one Phonological Phrases. The effect of Syntax is less clear. It is not constant and several syntactic distinctions (e.g. the complement-adjunct distinction) did not seem to have a corresponding effect on phrasing. Furthermore, one could always argue that any syntactic effect is in fact semantic in nature, and syntactic constituency and phonological phrasing is actually a reflection of a predicate-argument structure. In any case, the syntactic elements which were found to be more relevant were indeed maximal projections and phrase heads, each type corresponding to central notions in the two syntactically based approaches tested, namely Truckenbrodt’s (1999) and Nespor & Vogel’s (1986) respectively. Between these two proposals the latter was shown to better predict the data. Tables 5.1, 5.2 and 5.3 show the results of the evaluation.

PBs PREDICTED CORRECTLY	PBs NOT PRED.	PBs PREDICTED INCOR.
2547	126	312

Table 5.1: Evaluation of Nespor and Vogel’s proposal

²With all these problems this approach entails for the complexity of the grammar and its integration within present models of the language faculty

PBs PREDICTED CORRECTLY	PBs NOT PRED.	PBs PREDICTED INCOR.
1271	1399	237

Table 5.2: Evaluation of Truckenbrodt's proposal (Clauses are not dealt with individually)

PBs PREDICTED CORRECTLY	PBs NOT PRED.	PBs PREDICTED INCOR.
2020	650	375

Table 5.3: Evaluation of Truckenbrodt's proposal (Clauses delimited with PP breaks)

Furthermore, Nespor & Vogel's (1986) proposal did not have to be adjusted so as to process each clause separately, as this resulted directly from their definition of the Phonological Phrase. For Truckenbrodt's (1999) proposal on the contrary, this adjustment proved crucial for improving the proposal's predictions. Additionally, by discarding the complement-adjunct distinction, the accuracy of Nespor & Vogel's (1986) proposal did not decrease and we did away with a distinction which is not so easy to draw, nor is it widely accepted to stand. Finally, given that the head is the most important, central element of the syntactic projection, it does not seem unreasonable that the 'head property' would be marked and made eminent through phrase finality.

There is, however, one issue that poses problems to their proposal, the 'optional restructuring' parameter. The complete table for the evaluation of Nespor & Vogel's (1986) is actually the following, which includes 621 instances of restructuring, which would otherwise be included in the 'incorrectly predicted' category.

PREDICTED CORRECTLY	NOT PRED.	PREDICTED INCORR.	RESTRUCTURING
2547	126	312	621

Table 5.4: Evaluation of Nespor and Vogel's proposal

Based on our data, there are no clear-cut conditions under which restructuring should apply. Length seems to be the most important parameter that determines its application, but there is still a considerable amount of variability with respect to it. To apply

restructuring to the test corpus, we assumed that restructuring occurs only when the following phonological phrase is a one or two-word sequence³. The results are shown in table 5.5. Even though the accuracy of the predictions decreases, the proposal still makes better predictions with respect to the other two.

PBs PREDICTED CORRECTLY	PBs NOT PRED.	PBs PREDICTED INCOR.
2427	246	540

Table 5.5: Evaluation of Nespor and Vogel's proposal (Applying restructuring)

So, to sum up, in order to predict prosodic phrasing, the following approach seems to be the most promising. For cases in which there is no clear effect of discourse structure (more like the default cases):

- Each clause is processed separately.
- The clause is then divided into Phonological Phrases according to Nespor & Vogel's (1986) definition of the PP domain, that is
A Phonological Phrase consists of a lexical head X and all constituents on the non-recursive side of the head's(X) phrase, which are contained in the same maximal projection.
This condition should be stated in terms of an *extended projection* (see Chapter 3) as far as noun projections are concerned, in order to conform with the latest developments in syntactic theory.
- Restructuring of the Phonological Phrase may occur with respect to:
 - the length of the following Phonological Phrase (e.g. Phonological Phrases consisting of one or two words should attach to the previous one)
 - the lexical properties of the head (e.g. restructuring should occur after all link verbs)⁴

³Restructuring was considered not to apply, in the case of apposition structures of any length

⁴Several other empirical observations may be used, in order to improve the theory's performance. Allowing, for example, restructuring in subject-verb sequences where the subject is a one-word phrase,

There are, however, more parameters that need to be tested and would contribute to a better understanding and more accurate predictions of prosodic phrasing. For example, a more in depth analysis of length with respect to balance as well (cf Gee & Grosjean (1983)) should prove very useful. That is, the effect of length should be assessed on a syntagmatic level as well. In several instances in the corpus the length of phonological phrases seemed to be a choice that was active with respect to the whole utterance or big parts of it. For example:

(Green was transferred)(to a low lead area).

(Her base salary)(remained the same)(but she lost)(incentive pay).

(As an example)(the two term)(District Attorney)(cites what he calls)(the criminal justice system's)(of total crisis).

(The policy also is humiliating)(says Joanne Leard,)(a union health and safety monitor)(who worked at Johnson Controls)(for sixteen years).

Furthermore, in the beginning of this paper, we mentioned the violability of the Non-Recursivity constraint. Allowing for recursion in Phonological Phrases, could explain a lot of cases in which Truckenbrodt's (1999) theory failed. For example the phrase break after *attached* in the next utterance could be accounted for in terms of a nested structure.

(The device)(is attached to a plastic wristband)

(The device)(is attached (to a plastic wristband))

However, the adoption of such recursive structures can only be based on theory external factors, justified by independent evidence, such as differences in boundary strength (as the ones observed in Ladd (1996)) or other acoustic cues.

In addition, the study could also be expanded to examine the influence of speech rate, include different speech styles and different speakers as well, so that any consistencies or inconsistencies with respect to these variables can be determined, along with the extent of each factor's influence, the order and the level at which each factor is 'active'.

would prevent a significant number of incorrect predictions. However, dealing with the type of phrase breaks which were not predicted by the theory turns out to be more difficult. These types involve mainly pre-nominal modifiers and they were shown to be the exception rather than the norm.

One final question that should be addressed is the extent that each theory's claims can be connected to practical applications, the possibility of their implementation and integration in speech synthesis systems.

Most current synthesizers make little use of syntax (often replaced by parts of speech tagging) and no actual use of discourse, since most of the times each utterance is processed separately. In the Festival speech synthesizer's manual, for example, the following criteria are suggested as a good first approach to predicting phrasing and are further used as predictors in stochastic models.

Punctuation A phrase break is inserted at every punctuation mark.

Content vs Function words A phrase break is inserted after a content word and before a functional one.

Number of words since previous and following punctuation There has to be at least a specific number of words (e.g. five) since the previous and before the following punctuation mark for a phrase break to be inserted. This way very short Phonological Phrases are avoided.

The following paragraph corresponds to the predicted phrasing according to the criteria proposed above:

(In a study of seventy-two ballot issues in Massachusetts and three other states), (Boston University political science professor), (Betty Zisk), (found that eighty-eight percent of the battles were won)(by the side that spent the most money)

Comparing it with the actual phrasing (shown below), one can conclude that while the algorithm does not make a false insertion, it nevertheless, underpredicts.

(In a study)(of seventy-two ballot issues)(in Massachusetts)
 (and three other states),(Boston University)(political science professor),(Betty Zisk),
 (found that eighty-eight percent of the battles were won)(by the side)(that spent the
 most money)

However, this is still a fairly good approximation, especially if one considers the reliability of syntactic parsers and the difficulty of obtaining the necessary semantic and discourse information. There have, for example, been attempts to integrate discourse information in a speech synthesiser, but they have mostly involved simple dialogic systems, in which Information Structure was easily and unambiguously identified (Prevost & Steedman 1996). The case does seem to be that in order to apply the constraints or criteria that each theory suggests, one would need to assume that the input to the speech synthesiser would be the output of a Machine Translation or any Language Generation system, where a syntactic parse tree or Information Structure marking would be readily available⁵. Nevertheless, apart from the theoretical interest that these proposals on the Syntax-Phonology interface and ultimately on human cognition do have, one can only expect that the information they require will ultimately become available for speech synthesis or any other applications, so that eventually Phrasing at least, could be predicted based on something more than part of speech tags.

⁵Among the three different proposals Nespor & Vogel's (1986) should be the easiest to implement, since it makes no use of discourse information and does not involve a complement-adjunct distinction which even for humans is difficult to make

Appendix A

Test Corpus

Following are the news stories which were used for the evaluation of the various proposals. They are all included in the bu_radio/f2b/radio/ directory. The numbers between words correspond to ToBI phrase break indices.

STORY 1

a 1 nineteen 1 eighteen 4 state 1 constitutional 1 amendment 4 ; 5 made 1 Massachusetts 4 one
1 of 2 twenty 1 three 1 states 4 ; 5 where 1 citizens 4 can 1 enact 1 laws 4 by 1 plebiscite 4 ; 6
the 1 first 2 ballot 1 initiative 4 adopted 1 by 1 Bay 1 Staters 4 ; 5 was 0 a 1 nineteen 1 twenty
1 measure 4 ; 5 that 1 defined 3 cider 2 and 1 beer 3 as 1 non-intoxicating 4 ; 5 exempting
1 them 4 from 1 prohibition 4 ; 6 a 1 nineteen 1 forty 1 eight 1 initiative 4 to 1 legalize 3
contraceptives 3 was 1 defeated 4 ; 5 but 1 perhaps 2 the 1 State's 2 most 1 famous 1 initiative
4 was 1 proposition 3 two 1 and 0 a 1 half 4 ; 5 one 1 In 0 a 1 series 4 of 1 recent 2 anti-tax 1
measures 4 on 1 ballots 2 across 2 the 1 country 4 ; 6 Since 1 nineteen 1 sixty 1 eight 4 ; 5 voter
1 initiated 4- propositions 4 have 1 mushroomed 4 by 1 more 1 than 2 six 1 hundred 3 percent
4 ; 6 Heightened 1 awareness 3 of 0 the 1 process 4 ; 5 may 1 be 1 one 1 reason 4 why 2 more
1 petitions 2 were 1 certified 2 this 1 year 1 In 1 Massachusetts 3 than 1 ever 1 before 4 ; 5 says
1 the 1 Secretary 1 of 1 State's 3 Chief 1 Legal 1 Counsel 4 David 1 Sullivan 4 ; 6 Director 3-
of 1 the 1 McCormick 1 Public 1 Affairs 1 Institute 3- at 1 U-Mass 1 Boston 4 ; 4a Doctor 1
Ed 1 Beard 4 ; 5b says 2 the 1 push 3 for 2 do 0 it 1 yourself 1 lawmaking 4 comes 1 from 1
voters 3- who 1 feel 1 frustrated 4 by 1 their 1 elected 2 officials 4 ; 6 but 1 the 1 initiative 1

process 4 worries 1 Beard 4 and 1 other 1 political 1 scientists 4 ; 5 who 1 say 1 it 1 can 0 offer 3 simplistic 1 solutions 4 for 1 complex 1 problems 4 such 1 as 1 taxation 3 and 1 nuclear 1 energy 4 ; 6 Yet 2 Beard 1 remains 4 ambivalent 4 ; 5 because 2 he 1 can't 1 totally 4 write 0 off 1 an 1- institution 4 ; 5 that 1 makes 2 democracy 4 more 1 participatory 4 ; 6 In 1 a 0 study 3 of 1 seventy 1 two 2 ballot 1 issues 4 In 1 Massachusetts 3 and 2 three 1 other 1 states 4 ; 5 Boston 1 University 3 political 1 science 1 professor 4 ; 4a Betty 1 Zisk 4 ; 5b found 1 that 2 eighty 1 eight 1 percent 2 of 0 the 1 battles 2 were 1 won 4 by 1 the 1 side 3 that 1 spent 2 the 1 most 2 money 4 ; 6 Still 4 Zisk 1 applauds 4- voter 1 initiatives 4 ; 5 saying 4- they 1 shouldn't 2- be 1 held 3 to 1 a 1 higher 1 standard 4 than 1 laws 3 enacted 2 by 1 legislatures 4 ; 6.

STORY 2

the 1 policy 1 also 1 is 2 humiliating 4 says 2 Joanne 1 Leard 4 ; 5 a 1 union 1 health 1 and 0 safety 1 monitor 4 ; 5 who's 1 worked 1- at 1 Johnson 1 Controls 4 for 1 sixteen 1 years 4 ; 6 and 2 Virginia 1 Green 1 says 3 in 1 her 1 case 4 the 1 policy 1 is 2 infuriating 4 ; 6 She'd 1 had 1 her 1 children 4 before 2 she 1 began 1 working 2 at 1 Johnson 1 Controls 4 ; 5 and 1 didn't 1 worry 3 about 2 lead 1 exposure 4 ; 6 in 1 fact 4 Green 1 says 2 she 1 enjoyed 1 her 1 job 4 as 1 a 1 plate 1 stacker 4 ; 6 but 0 when 1 the 1 fetal 1 Protection 1 policy 3 took 1 effect 4 ; 5 Green 2 was 1 transferred 3 to 1 a 1 low 1 lead 2 area 4 ; 6 her 1 base 1 salary 3 remained 1 the 0 same 4 ; 5 but 1 she 1 lost 3- incentive 1 pay 4 ; 6 Green 2 was 3 fifty 1 years 1 old 3 at 1 the 1 time 4 ; 6 Green 4 several 1 of 1 her 1 Bennington 1 co-workers 4 and 1 employees 3 at 1 some 1- of 1 Johnson 1 Controls' 3- fifteen 1 other 1 manufacturing 1 plants 4 filed 1 grievances 4 ; 6 which 1 led 3 to 1 a 2 class 1 action 2- lawsuit 4 ; 6 Company 1 officials 3 refused 1 to 1- comment 4 ; 5 but 1 their 1 position 1 paper 4 on 1 the 1 fetal 1 Protection 1 policy 4 ; 5 maintains 3 that 1 certain 2 blood 1 lead 1 levels 4 while 2 safe 1 for 1 adults 4 ; 5 cause 2 birth 1 defects 4 If 1 passed 2 through 1 the 1 placenta 3 to 1 a 1 fetus 4 ; 6 According 1 to 0 the 1 document 4 ; 5 it 1 would 1 be 1 morally 2 reprehensible 4 ; 5 to 1 knowingly 2 poison 4 unborn 1 children 4 ; 5 by 1 allowing 3- fertile 1 women 4- to 1- work 4 in 1 high 1 lead 1 areas 4 ; 6 Doctor 2 Jay 1 Julian 1 Chism 4 ; 5 Director 2 of 0 the 1 Kennedy 1 Institute's 2 lead 1 program 3 at 0 Johns 1 Hopkins 1 University 4 ; 5 has 1 studied 2 lead 1 poisoning 4 for 1 some 1 thirty 1 five 1 years 4 ; 6 in 1 the 1 case 1 against 1 Johnson 1 Controls 4 ; 5 Chism 1 provided 3- an 1 affidavit 4 on 1 the 1 company's 3- behalf 4 ; 6 He 1 says 3 unless 2 lead 1 standards 2 are 1 lowered 4 throughout 1 the 1 workplace 4 ; 5 the 1 only 1 way 1 to 1 ensure 3 that 1 fetuses 4- are 1 not 1 at 1 risk 4 ; 5 is 1 to 1- remove 3 from 2 lead 1 areas 4 ; 5 anyone 3 who 1 may 1 become 1

pregnant 4 ; 6 but 2 research 2 on 1 pregnancy 2 and 0 lead 1 poisoning 4 is 1 not 1 conclusive 4 According 2 to 1- University 1 of 1- Wisconsin 2 Labor 1 Law 1 Professor 4 Karen 1 Klaus 4 ; 6 There's 1 also 1 evidence 3 she 1 says 4 ; 5 that 1 men 3 with 1 high 2 blood 1 lead 1 levels 4 may 4- through 1 their 1 sperm 4- endanger 1 their 1 offspring 4 ; 6 Klaus 1 represented 4- the 1 U 4 in 1 its 1 appeal 4 ; 6 she 1 says 2 similar 3 fetal 1 Protection 1 policies 4 have 1 been 1 struck 1 down 3 by 1 other 1 federal 1 courts 4 ; 5 on 1 the 1 grounds 3 that 1 they're 2 discriminatory 4 ; 6 but 1 times 2 have 1- changed 4 she 1 says 4 ; 5 because 1 of 1- the 1 nationwide 1 debate 3 over 1 abortion 4 and 2 fetal 2 rights 4 ; 6.

STORY 3

This 1 is 1 Margo 1 Melnicove 4 At 0 a 1 news 1 conference 1 yesterday 4 in 1 Mission 1 Hill 4 near 1 the 1 site 1 of 1 the 2 Stuart 1 shootings 4 ; 5 Boston 1 mayor 2 Ray 1 Flynn 3 vehemently 1 denied 1 charges 4 that 1 there's 1 a 1 double 1 standard 1 in 1- the 1 City 4 ; 5 that 2 white 1 life 4 is 1 more 1 important 4 than 2 black 1 life 4 ; 6 the 1 problem 3 ; 3a said 1 Flynn 4 ; 4b is 1 that 1- there 1 aren't 0 enough 2 financial 3 and 2 human 1 resources 4 ; 5 to 1- protect 1 the 1 public 1 safety 4 and 1 improve 1 the 1 quality 2 of 1 life 4 ; 6 Federal 2 and 1 state 2 budget 1 cuts 4 have 1 created 2 a 1 public 1 service 1 crisis 4 he 1 said 4 ; 5 with 2 drug 1 treatment 4 and 1 law 1 enforcement 4 only 2 two 1 of 1- the 1 casualties 4 ; 6 Flynn 2 offered 1 a 1 remedy 4 that 1 he 1 admits 3 is 1 not 1 likely 3 to 0 be 1 popular 4 ; 6 Flynn 1 said 3 with 2 local 1 option 1 taxes 4 such 1 as 1- an 1 increase 2 in 1 sales 1 tax 4 ; 5 the 1 money 1 goes 4 to 1- the 1 community 4 that 1 generates 1 it 4 ; 6 For 1 example 4 a 1 one 1 cent 2 sales 1 tax 1 increase 2 said 1 Flynn 4 ; 5 would 1 add 2 thirty 1 five 1 million 1 dollars 2 a 1 year 4 to 1- the 1 city's 2 treasury 4 ; 6 the 1 mayor 2- said 3 soon 2 he'll 1 seek 2 legislative 1 approval 3 of 1 such 1- a 1 tax 4 ; 5 with 1 Boston's 1 share 3 earmarked 4 For 2 public 1 safety 4 ; 5 youth 1 outreach 1 programs 4 ; 5 and 1 educational 1 reform 4 ; 6 yesterday 3 the 1 Boston 1 City 1 Council 3 called 2 For 1- an 1 emergency 1 meeting 2 with 1 Flynn 4 next 1 Monday 4 ; 5 to 1- put 1 other 1 strategies 3- on 1 the 1 table 4 ; 5 including 2- a 1 new 1 court 4 to 1 handle 2 drug 1 related 1 cases 4 ; 5 expansion 1 of 1 neighborhood 3- crime 1 watch 1 efforts 4 ; 5 and 1 a 1 beefed 1- up 1 budget 3 For 1 police 4 ; 6 These 1 ideas 3 have 1 been 2 floated 1 before 4 ; 5 Louis 1 Alisa 4 President 2 of 1- the 1 N 1 double 1 ACP's 2 Boston 1 branch 4 says 1 it's 1 sad 4 that 0 it 1 takes 2 the 1 death 3 of 1 a 1 white 1 woman 4 from 1 the 1 suburbs 4 ; 5 to 1 create 2 an 1 urgency 4 to 1- rid 1 Boston 4- of 1 drugs 3 poverty 3 and 1- guns 4 ; 6 But 1 the 1 Stuart 2 tragedy 3 may 1 send 2 another 1 message 4 says 2 the 1 head 1

of 1 Boston 1 University's 3 Urban 1 Studies 2 and 1- public 1 Policy 2- Program 4 Joseph 1 Boscan 4 ; 6 Boston 2 is 1 already 2 divided 4 ; 4a says 1 Boscan 4 ; 4b in 1 terms 1 of 1 class 3 race 3 and 1 ethnicity 4 ; 6 and 1 in 1- a 2 fragmented 1 City 4 he 1 says 4 ; 5 one 1 finds 2 a 1 great 1 deal 3- of 1 violence 4 ; 5 For 1 WBUR 4 ; 5 I'm 1 Margo 1 Melnicove 4 ; 6

STORY 4

This 1 woman 1 receives 3 three 1 hundred 1 dollars 2 a 1 month 4 under 2 General 1 Relief 4 plus 2 four 1 hundred 1 dollars 1 a 1 month 4 in 1 A 1 benefits 4- for 1 her 1 son 4 who 1 is 1 a 1 U 2 citizen 4 ; 6 She's 1 among 3- an 1 estimated 3 five 1 hundred 4 illegal 1 aliens 4 on 2 General 1 Relief 4 ; 5 out 1- of 1 the 1 state's 2 total 3 illegal 2 immigrant 1 population 4- of 1 one 1 hundred 1 thousand 4 ; 6 General 1 Relief 3 is 1 for 2 needy 1 families 4 and 2 unemployable 1 adults 4 ; 5 who 1 don't 2 qualify 3 for 1 other 2 public 1 assistance 4 ; 6 Welfare 1 Department 1 spokeswoman 4 Michael 1 Reganburg 4 ; 5 says 2 the 1 state 2 will 1 save 3 about 2 one 1 million 1 dollars 2 a 1 year 4 if 2 illegal 1 aliens 4 are 1 denied 4 General 2 Relief 4 ; 6 Reganburg 1 says 4- if 2 state 1 finances 1 improve 4 benefits 2 may 1 become 1 available 4 to 2 illegal 1 aliens 3 again 4 ; 6 but 1 several 1 legislators 4- have 1 filed 1 bills 4 to 1- make 1 sure 2 that 1 doesn't 1 happen 4 ; 6 Republican 1 Senator 1 David 1 Lock's 1 bill 4 ; 5 would 2- permanently 1 deny 4 illegal 1 aliens 4 all 1 state 1 benefits 4 except 2 emergency 1 medical 3 assistance 4 ; 6 but 0 Advocates 4- for 2 undocumented 1 immigrants 4 ; 5 say 2 only 1 the 1 most 2 desperate 4- apply 4 for 2 public 1 assistance 4 ; 5 and 1 without 1 it 4 they'll 1 be 1 hungry 4 and 1 homeless 4 ; 6 Advocates 1 also 1 say 4- most 1 immigrants 3- come 1 here 3- for 2 legitimate 1 reasons 4 ; 5 but 1 they 1 can't 1 get 1 asylum 4 because 2 federal 2- immigration 1 policy 4- is 1 biased 4 ; 5 especially 2 against 2 Central 1 Americans 4 ; 6 Attorney 1 Michael 1 Altman 4 advises 3 the 1 Massachusetts 1 immigrant 3- and 1 Refugee 3 Advocacy 1 Coalition 4 ; 5 Altman 1 says 3 it's 1 immoral 4 to 1 deny 2 anyone 4 documented 1- or 1- not 4 life's 2 basic 1 necessities 4 ; 5 and 1- it 1- also 2 may 1 be 4 illegal 4 ; 6 Regulations 3 banning 4 General 1 Relief 4- for 2- illegal 1 aliens 4 take 1 effect 3 next 1 month 4 ; 5 Whether 1 or 1 not 3 the 1 ban 3 becomes 2 permanent 4 under 1 a 1 new 1 law 4 ; 5 Altman 1 says 4 the 1 change 4 will 1 be 1 challenged 4 in 1 court 4 ; 6 for 2- WBUR 4 I'm 1- Margo 1 Melnicove 4 ; 6

STORY 5

the 1 State 3- public 1 health 1 department 4 must 1 inspect 1 everything 3 from 1 milk 1 plants 4 to 1- hospitals 4 to 1 police 1 station 1 lockups 4 ; 6 State 1 inspectors 3 also 1 assist 4 local 1 boards 3- of 1- health 4 ; 6 These 1 part 1 time 1 boards 4 are 1 the 1 only 1 health 1 inspectors 4 in 1 one 1 hundred 0 and 1- sixty 1 two 4 small 2 Massachusetts 1 communities 4 ; 5 so 2 State 1 workers 1 often 1 step 1 in 4 ; 5 to 1 survey 2 restaurants 4 septic 1 systems 4 housing 4 ; 6 and 1 towns 3 that 1- employ 2 fulltime 4- public 1 health 1 agents 4 ; 5 often 1 turn 1 to 1- the 1 State 4 for 1 training 4 ; 6 State 1 public 1 health 2 Deputy 1 Commissioner 4 ; 4a Van 1 Dunn 4 ; 5b admits 2 recent 1 layoffs 4 will 1 mean 2 less 1 help 4 for 2 local 1 boards 4 ; 5 and 0 may 1- make 2 the 1 State 3 more 1 reactive 4 than 2 prevention 1 oriented 4 ; 6 Dunn 1 says 2 the 1 food 1- and 1- drug 1 Division 3 ; 3a for 1 example 4 ; 5b expects 1 its 2- annual 1 inspections 3 to 1- drop 4 by 1 fifteen 1 hundred 4 ; 6 But 1 Dunn 1 says 4- with 1 some 2 staff 1 reshuffling 4 he'll 1 make 2- the 1 best 4 of 1 a 1 bad 1 situation 4 ; 6 and 1 continue 1 to 1 protect 3 public 1 health 4 ; 6 this 1 sounds 3 like 1 a 2- physical 1 impossibility 4 to 1 greater 1 Boston 1 legal 1 services 2 housing 1 attorney 4 ; 4a Jeff 1 Purcell 4 ; 5b who 1 says 3 near 2- elimination 4- of 2 State 1 sanitarians 4 ; 6 portends 3 a 2 tenant's 1 rights 1 crisis 4 ; 6 the 1 State 1 used 1- to 1 employ 4- seven 1 sanitarians 4 now 2 two 3 must 1 cover 2 all 1 of 1 Massachusetts 4 ; 6 Their 1 work 2 is 1 cut 1- out 2 for 1- them 4 ; 5 with 1 inspections 3 of 1 prisons 4 State 1 operated 1 camp 1 grounds 2 and 1- swimming 1 pools 4 ; 5 and 1 farm 1 labor 1 camps 4 required 2 by 1 law 4 ; 6 Purcell 1 says 4- there's 2 no 1 way 3 they 1 can 1 also 1 take 1 on 4 ; 5 the 1 sanitarian's 2 traditional 1 role 4 of 2 inspector 1 of 1 last 1 resort 4 ; 5 when 1 local 1 officials 2 fail 4 to 2 crack 1 down 2 on 1 slum 1 lords 4 ; 6 Purcell 1 warns 4- that 1- this 1 winter 4 more 1 tenants 1 than 1 usual 4 will 1 have 0 to 1- put 0 up 1 with 2- inadequate 1 heat 2 and 1 hot 1 water 4 ; 5 while 1 they 1 await 2- a 1 State 1 sanitarian 4 who 2 may 1 never 1 show 4 ; 6 Director 1 of 0 the 0 Division 2 of 1 Healthcare 1 Quality 4 ; 4a Priscilla 1 Plato 4 ; 5b says 2 delays 3 are 1 also 1 expected 4 ; 5 in 1 Re-licensure 1 inspections 4 at 1 rest 1 homes 4 hospitals 4- and 1- clinics 4 ; 6 Re-licensure 1 inspections 4 used 1 to 1 occur 3 every 3- two 1 years 4 But 1 now 3 they 1 will 1 happen 3- every 1 four 3- or 1 six 1 years 4 ; 4a says 1 Plato 4 ; 5b whose 1 Division 3- this 1 week 3 lost 2 twelve 1 percent 4 of 1 its 1 inspection 4- work 1 force 4 ; 6 for 2 WBUR 4 I'm 1 Margo 1 Melnicove 4 ; 6 STORY 6

ice 2 is 1 a 1 smokable 4 potent 1 form 4- of 1 methamphetamine 4 or 1 speed 4 ; 6 a 1 few 1 hits 3 bring 1- on 1 a 1 high 4 that 1 lasts 2 For 1 hours 4 ; 6 ice 1 wreaks 1 havoc 2 on 1 the 1 body 4 causing 2 increased 1 blood 1 pressure 4 tremors 4 acne 4- and 1- other 2 physical

1 damage 4 ; 6 But 1 Doctor 1 Alonso 1 Plow 4 ; 4a Boston 1 City 1 Hospital's 4- Deputy 1 Commissioner 3- For 1 Public 1 Health 4 ; 6b says 3 what's 1 most 1 alarming 4 ; 5 are 1 the 2 psychological 3- effects 2 of 2 ice 4 once 1 the 1 initial 1 euphoria 4 subsides 4 ; 6 Plow 1 says 3 in 2 San 1 Diego 1 California 4 nearly 2 one 1 third 3- of 1 all 2 homicides 4 now 1 involve 1 people 3 high 1 on 1 ice 4 ; 6 But 1- by 2- the 1 time 3- the 1 drug 2 shows 1- up 0 in 1 Boston 4 ; 5 Plow 1 hopes 3 the 1 word 1 on 1- the 1 street 3 will 1 be 3- that 1 playing 2 with 1 ice 4 is 1- in 1 fact 4 playing 2 with 1 fire 4 ; 6 the 1 hospital 4 the 1 Mayor's 1 Office 4 the 1 school 1 system 4- Police 3 and 1 others 4 ; 5 are 1 working 1 on 1- an 1 educational 1 campaign 4 about 2 the 1 dangers 1 of 1 ice 4 ; 6 No 1 one 1 knows 2 how 1 much 1 time 2 they'll 1 have 4 to 1- warn 2 potential 1 users 4 ; 6 Los 1 Angeles 3 Police 1 detective 4- Milt 1 Dodge 4 ; 5 says 3 ice 1 is 1 beginning 3- to 1- surface 2- in 1 his 1 City 4 ; 5 with 1- a 1 couple 1 of 2 small 1 seizures 4- made 2 the 1 last 2 few 2 days 4 ; 6 Dodge 1 says 2 prevention 1 efforts 3 in 1 L 4 run 1 the 1 gamut 4 instead 1- of 1 focusing 3 on 1- a 1 particular 1 drug 4 ; 6 in 1 Chicago 4 a 1 Police 1 spokeswoman 2 says 4 ; 5 there's 1 fear 2 that 2 publicity 3 will 1 only 2 arouse 2 demand 2 For 1 ice 4 ; 5 so 1 until 1 the 1 drug 4 hits 1 Chicago 4 there'll 1- be 1 No 1 campaign 2 against 1 it 4 ; 6 But 1- in 1 Boston 4 Deputy 1 Superintendent 3- of 1 the 1 drug 1 Control 1 Unit 4 ; 4a James 1 Wood 4 ; 5b says 3 information 4 is 1 the 1 best 1 weapon 4 ; 6 Wood 1 applauds 4- the 1 upcoming 1 campaign 4 which 1 will 1 be 1 similar 4 to 1 a 1 media 3- and 2- classroom 1 blitz 4 waged 2 a 1 few 1 years 1 ago 4 ; 6 after 1 New 1 York 1 Police 4 tipped 1 off 1 Boston 4 ; 5 that 2 smokable 1 cocaine 4 or 1 crack 4 was 1 headed 2 north 4 ; 6 Wood 1 credits 3 the 1 city's 3- anti-crack 2 campaign 4 with 1 keeping 1 that 1 drug 4 under 3- Control 4 ; 6 Boston 1 has 2- a 2 crack 1 problem 4- ; 4a he 1 admits 4 ; 5b But 1- it's 1 nothing 4 compared 3- to 1 the 1 epidemic 4- in 1 other 1 cities 4 ; 6 and 1 maybe 3 says 1 Wood 4 ; 5 similar 3- education 1 efforts 4 ; 5 can 1 keep 1 a 1 lid 1 on 1 demand 4 For 1 the 1 latest 1 designer 3- drug 4 ice 4 ; 6 For 1 WBUR 4 I'm 1 Margo 1 Melnicove 4 ; 6

STORY 7

Boston 2- city 1 counselor 3- David 1 Scondras 4 who's 2- gay 4 is 2 plenty 2 angry 4 ; 5 but 1 not 2 at 1 Barney 1 Frank 4 who 1 paid 2 for 1 sex 4 when 1 he 1 was 1 still 3- in 1 the 1 closet 4 ; 6 Scondras 2 blames 4- Frank's 1 misconduct 4 ; 5 on 1- a 1 homophobic 1 society 4 that 2 forces 4 most 1 gays 4 ; 5 to 1 live 3 secret 1 lives 4 ; 6 So 1- why 1- hasn't 1 the 1 gay 2 community 4- rallied 1 around 1 Frank 4 with 1- a 1 march 4 or 2 other 1 show 4- of 1 widespread 1 support 4 ; 6 Bay 1 Windows' 2 Epperly 4- says 4 people 1 are 1 waiting 4 to 1-

see 2- if 1 the 1- debate 1 escalates 4 ; 5 into 1 a 1 overt 2 attack 4 against 2 gays 1 in 1 general 4 ; 6 David 1 Scondras 3- says 4 ; 5 many 1 gays 3 are 1 still 1 afraid 4 to 1- come 1- out 1- of 1 the 1 closet 4 to 1 take 2 a 1 public 2 stand 4 ; 6 but 1 this 1 customer 3- at 1 Fritz's 4 has 1 another 1 theory 4 ; 6

STORY 8

Barney 1 Frank's 1 district 4 stretches 3 from 2 upscale 2- Democratic 2- Newton 4 ; 5 through 1 affluent 3 mostly 1 Republican 1 suburbs 4 such 1 as 2 Wellseley 1- and 0 Dover 4 ; 5 to 1 the 1 working 1 class 1 milltown 4- of 1 Fall 1 River 4 ; 6 his 2- diverse 2- constituents 4 have 1 agreed 2 repeatedly 4 that 1 Frank 3 serves 1 them 1 well 4 ; 5 as 1- a 1 champion 3 of 2- equal 1 rights 4 affordable 1 housing 4 the 1 elderly 4 and 1 reduced 2- defense 1 spending 4 ; 6 Frank 1 won 2 his 2 fifth 2 congressional 1 term 3- in 1 nineteen 1 eighty 0 eight 4 ; 5 with 1 seventy 2 percent 3- of 1 the 1 vote 4 ; 5 despite 1 his 1 public 1 announcement 4- a 1 year 1 earlier 4 that 1 he's 0 a 1 homosexual 4 ; 6 Frank's 2 more 1 recent 1 admission 4 ; 5 that 1 he 1 paid 2 For 1 sex 4 before 1 he 2 came 1 out 0 of 1 the 1 closet 4 ; 5 and 0 that 0 he 1 once 2 employed 3 a 1 male 1 prostitute 4 as 1 his 1 housekeeper 0 and 1- driver 4 ; 5 initially 2 angered 4 many 1 constituents 4 according 2 to 1 Dorothy 1 Wrikhard 4 ; 6 But 1 the 1 director 3- of 0 Frank's 2 district 1 offices 4 says 3- now 4 calls 2 of 1 support 4 are 2 pouring 1 in 4 ; 6 Wrikhard 1 says 2 supporters 4 have 1 outnumbered 1 detractors 4 by 1 about 2 eight 1 to 1 one 4 ; 5 and 1- the 1 Frank 1 campaign 3 has 1 received 2 seventeen 1 hundred 1 dollars 4 in 1 unsolicited 2 donations 4 ; 5 with 1- another 2 three 1 thousand 1 dollars 2 pledged 4 ; 6 But 1- a 1 survey 4- in 1 Frank's 1 district 4 last 1 week 4 ; 5 showed 1 a 1 precipitous 2 fourteen 1 point 1 drop 4 in 1 his 2 favorability 1 rating 4 ; 5 and 1 pollster 2 Jerry 1 Travinski 3- says 4 ; 5 the 1 numbers 3 may 1 not 1 have 2 bottomed 1 out 4 ; 6 Some 1 pollsters 1 say 4 voters 4 most 1 likely 3- to 0 be 1 turned 1 off 4 by 1 Frank's 1 conduct 4 ; 5 are 1 newcomers 2 to 1 the 1 district 4 who 1 are 1 unaware 3 of 1 his 2 political 1 record 4 ; 5 and 1 conservative 3 blue 1 collar 3 ethnic 1 Democrats 4 ; 5 who 1 may 1 have 1 tolerated 3 Frank's 1 homosexuality 3- before 4 ; 5 But 1 now 4- are 1 disturbed 4- by 1- a 1 male 1 prostitute's 3- claim 4 ; 5 that 2 he 1 conducted 1 business 4 out 0 of 1 Frank's 2 Washington 1 apartment 4 ; 6 Several 1 constituents 4 are 1 planning 2 petition 1 drives 4 on 1 Frank's 1 behalf 4 ; 5 saying 3 he 1 still 1 can 1 be 4 an 1 effective 1 representative 4 ; 6 But 1 Some 1 political 1 analysts 3- say 4 Frank's 1 credibility 3 in 2 Congress 3 is 1 shot 4 ; 6 and 2 Some 1 Democratic 1 insiders 4 afraid 4- of 1 how 1 a 1 prolonged 3- Ethics 1 Committee 3- Investigation 4 might 1 affect 2

the 1 party 4 ; 5 seem 1 to 1 be 1 increasing 2 the 1 pressure 2 on 1 Frank 3 to 1 resign 4 ; 6
 But 1 Frank 1 says 4 no 1 one 1 in 1 the 1 leadership 4 has 2- asked 1 him 3 to 1 step 1 down 4
 ; 6 For 1 WBUR 4 I'm 1 Margo 1 Melnicove 4 ; 6

STORY 9

During 1 the 1 past 3 one 1 hundred 2 and 1 fifty 1 years 4 ; 5 seventy 1 two 1 species 3 in 1
 Massachusetts 4 have 1 become 2 extinct 4 ; 6 Initially 4 they 1 were 1 ravaged 2 by 1 hunters
 4 ; 5 later 3 by 1 pollution 4 ; 5 and 1 most 1 recently 4 by 2- real 1 estate 3- development 4 ;
 6 Today 4 state 1 biologists 3 consider 2 four 1 hundred 1 and 0 three 4 native 1 plant 2 and 1
 animal 1 species 4 ; 5 to 0 be 1 of 1 special 1 concern 4 threatened 4 or 1 endangered 4 ; 6 Only
 1 twenty 1 two 3 are 1 protected 3 by 2- federal 1 law 4 ; 6 It's 1 up 1 to 1 each 1 state 4 to 1
 protect 2 other 1 rare 1 species 4 found 2 within 1 its 1 borders 4 ; 6 at 1 least 2 thirty 1 states 3
 have 1 such 1 laws 4 ; 5 Massachusetts 4 does 1 not 4 ; 6 Craig 1 Mactenel 4 is 1 legislative 1
 liaison 4 for 1- the 1 Fisheries 2 and 1 Wildlife 3 Department 4 ; 6 the 1 bill 3 would 1 impose
 2 heavy 1 fines 4 for 2- killing 3 or 1 selling 2 any 2 of 1 the 1 state's 2 rare 1 species 4 ; 5 and
 1 even 2 stiffer 1 penalties 4 for 1 altering 2 the 1 habitat 4 of 1 threatened 3- or 1 endangered
 1 species 4 without 1 a 1 special 1 permit 4 ; 6 Garon 1 Bresnick 4 ; 4a General 1 Counsel 3 of
 1 the 1 Massachusetts 2 Homebuilders 1 Association 4 ; 4b says 2 the 1 effect 3- would 1 be
 4- to 1 endanger 3 developable 1 land 4 ; 6 he 1- says 3- property 1 owners 4 already 2 have 1
 their 1 hands 1 tied 4 ; 5 by 0 a 1- state 1 law 3 protecting 1 wetlands 4 ; 6 and 1- now 3 the 1
 state 4 wants 1 to 1 encroach 4 on 1 building 1 sites 4 that 0 are 1 high 0 and 1- dry 4 ; 6 But
 1 Henry 1 Willsey 4 ; 4a head 0 of 1 the 1 state's 2 Natural 1 Heritage 3 and 1 endangered 1
 species 1 Program 4 ; 4b says 2 the 1 measure 3 would 1 affect 2 relatively 3- few 1 acres 4
 perhaps 1 Only 3- seventy 1 five 1 hundred 3 throughout 1 the 1 state 4 ; 6 Willsey's 1 Program
 4 has 1 spent 1 the 1 last 1 decade 4 ; 5 collecting 1 data 3 on 1 rare 1 species 4 ; 5 and 1 he
 1 knows 2 where 1 they 1 are 4 ; 5 Cape 1 Cod 3- and 0 the 1 islands 4 the 1 South 1 Shore
 3 near 1 Plymouth 4 and 1 parts 1 of 1 the 1 Berkshires 4 ; 5 to 0 name 1 a 1 few 3 critical 2
 habitat 1 areas 4 ; 6 But 1 Willsey 1 says 4 he 1 won't 1 publicize 4 specific 1 sites 4 ; 5 until
 1- after 3 the 1 endangered 1 species 1 bill 3 becomes 1 law 4 ; 6 the 1 issue 3 ; 3a he 1 says 4
 ; 4b should 1 not 1 be 1 influenced 4 by 2 politics 4 and 2 who 1 owns 3 what 1 property 4 ; 5
 But 1 should 1 be 1 based 4 on 1 biology 4 ; 6

Critics 3- of 1 the 1 proposed 1 law 4 1 say 1 the 1 state 2 can't 1 afford 0 it 4 in 1 these 2

financially 1 troubled 1 times 4;5 But 1 Willsey 1 says 4 the 1 bill 3 would 1 be 2 mostly 1 funded 4 by 2 voluntary 1 contributions 4- on 1 state 1 income 1 tax 1 forms 4;6 and 0 he 1 says 3 there's 2 ten 1 million 2 dollars 4- from 1 bond 1 sales 4 for 1 habitat 1 acquisition 2 on 1 hand 4;5 should 1 the 1 state 3 have 0 to 1 buy 1 out 4- some 1 land 1 owners 4;6 But 1 developers 3- say 4 land 1 takings 4 under 1 the 1 endangered 1 species 1 bill 4 would 1 be 3 inequitable 4 If 1 not 2 illegal 4;5 and 1 the 1 state 3- may 1 be 1 hit 3 with 1 a 1 series 4 of 1 costly 1 lawsuits 4;6 the 1 measure's 3 price 1 tag 4 will 1 be 1 reviewed 4 by 1 the 1 House 2 Ways 0 and 1 Means 1 committee 4;6 Greenfield 1 representative 3 Carmen 1 Bule 4-4a a 0 member 2 of 1 the 1 committee 3 and 1 sponsor 2 of 0 the 1 bill 4;4b is 1 hoping 3 for 2 timely 3 consideration 4;6 If 1 the 1 measure 3- wins 3 the 1 legislature's 2 final 1 approval 3 before 1 this 1 session 3- ends 4 ; 5 it 0 will 1 take 1 at 1 least 1- a 1 year 4 to 1 finalize 3- regulations 4 ; 5 and 0 to 1 publicly 1 outline 2 the 1 habitat 4 of 1 the 1 state's 2 most 1 endangered 3 species 4 ; 6 for 2 WBUR 4 I'm 1- Margo 1 Melnicove 4 ; 6

STORY 10

the 1 contract 4- was 1 unveiled 2 last 2- May 4 after 2 eight 1 months 3 of 1 negotiations 4 and 1 hailed 2 by 1 educators 4- as 2 historic 4 ; 6 the 1 pact 4- between 1 the 1 teacher's 1 union 3 and 0 the 1 school 1 Department 4 ; 5 calls 3- for 1 an 2- unprecedented 3 shift 1 of 1 power 4 from 2 central 1 office 1 administrators 4 to 1 those 3 who 1- know 3 each 1 school 1 best 4 ; 5 teachers 4 parents 4 and 1 students 4 ; 6 it's 1 called 4 school 1 Based 4- Management 4 ; 5 and 1 Boston 1 teachers 1 union 1 President 4 ; 4a Ed 1 Dougherty 4 ; 5b touts 1 it 2 as 1 essential 4 to 1 turning 1 around 4- the 1 city's 3 troubled 2 school 1 system 4 ; 6 When 1 the 1 contract 3 was 1 announced 4 Boston 1 mayor 2 Ray 1 Flynn 4 praised 2 the 1 reforms 2 it 1 promises 4 ; 5 but 0 in 1- the 1 next 1 breath 4 he 1 cautioned 4 that 1 the 1 city 2 would 1 be 2 hard 1 pressed 3 to 1 pay 4 for 1- the 1 eighty 1 nine 2 million 1 dollar 1 package 4 ; 6 Since 1 then 4- the 1- city's 1 financial 1 outlook 4 has 1 become 2 even 1 gloomier 4 ; 6 in 1 an 1 effort 2 to 1 avert 2 today's 1 walkout 4 Flynn 1 offered 0 to 0 fund 3 a 2- one 1 year 1 contract 4 ; 5 but 1 that 1 would 1 require 4- the 1 union 4- to 1 negotiate 4- a 1 whole 1 new 1- agreement 4 ; 6 Ed 1- Dougherty 3- says 4 that's 1 impossible 4 ; 6 Some 1 people 3- blame 4- the 1 Dukakis 1 administration 4 for 1 it's 3 budget 1 bungling 4- and 2 cuts 1- in 2- local 1 aid 4 ; 5 others 4- the 1 legislature 4 for 0 its 2 failure 4 to 1 approve 3 new 1 taxes 4 ; 6 but 1 many 1 teachers 1 and 1- parents 4 blame 1 mayor 1 Flynn 4 ; 5 saying 3 he 1 should 0 have 1 worked 1 harder 4 on 1 behalf 2 of 1 the 1 contract 4 instead 1 of 1 politicizing 4 for 0 an 1 appointed 1 school

1 committee 4 ; 6 Some 1 teachers 3 theorize 4 that 1 Flynn's 3 trying 0 to 0 get 1 back 1- at 1 the 1 union 4 ; 5 for 1 having 1 opposed 2 his 1 referendum 4 on 2- school 1 committee 1 restructuring 4 ; 6 but 1 Dougherty 4- has 1 another 1 theory 4 ; 6 he 1 says 3- the 1 mayor 4 just 2 doesn't 1 want 3- the 1 contract 3- funded 4 ; 6 Jeff 1 Conley 4 heads 2 the 1 Boston 2 Finance 3- Commission 4 ; 5 a 1 watchdog 1 agency 4 created 3- by 1 the 1 city 1 charter 4 ; 6 the 1 union 1 estimates 4 that 0 at 0 least 2 thirty 1 million 1 dollars 2 are 1 available 4 in 1 unspent 1 money 1 from 1 last 1 year 4 ; 6 but 2 Conley 1 says 4 that 0 amounts 2 to 0 only 2 ten 1 million 1 dollars 4 and 2- several 1 budget 1 items 4 are 2- vying 2 for 0 the 1 funds 4 ; 6 the 1 mayor's 2- education 1 Advisor 4 ; 4a Ellen 1 Guiny 4 ; 5b says 0 it 1 would 1 be 3- irresponsible 3 for 1 Flynn 4- to 0 commit 3 to 0 the 1 three 1 year 1 contract 4 given 2 current 1 revenue 1 forecasts 4 ; 6 but 0 she 1- says 3 those 1 who 1 criticize 2 the 1 mayor 4 for 2 backing 1 down 4- on 1 his 1 commitment 3- to 0 education 4 ; 5 are 1 wrong 4 ; 6 Guiny 1 says 3 the 1 mayor 1 intends 3 to 1 continue 1 to 1 push 2 for 1 new 1 taxes 4 ; 5 and 2 perhaps 4- within 0 a 1 month 1- or 1 two 4 the 1 political 1 winds 3 will 2- shift 0 enough 1- in 1 his 1 direction 4 to 1 allow 1 him 4- to 1 support 4- the 1 teacher's 1 contract 4 ; 6 the 1 Union's 3- Ed 1 Dougherty 3- says 4 ; 5 he 1 won't 1 know 3- until 1 after 3- a 1 board 1 meeting 3 next 1 week 4 just 2 how 1 long 3- the 0 union 2 can 1 wait 4 ; 5 before 2 calling 3 for 0 a 2 full 1 scale 1 strike 4 ; 6 for 2- WBUR 4 I'm 1- Margo 1 Melnicove 4 ; 6

STORY 11

the 1 Teen 1 Health 1 Center 4- at 1 Cambridge 2 Ringe 1 and 1 Latin 4 ; 5 is 1 visited 3- about 2 one 1 thousand 1 times 3 each 1 month 4 ; 5 by 1 students 3 seeking 1 everything 4 from 1 aspirin 3 to 1 immunizations 4 to 1 prenatal 1 care 4 ; 6 Within 0 a 1 few 1 weeks 4 condoms 3 will 1 be 1 added 2 to 1- the 1 list 4 ; 6 It 1 hasn't 1 been 1- determined 3 exactly 2 who 1 will 1 dispense 1 them 4 or 2 how 1 many 3- a 1 student 3 will 1 be 1 allowed 2 to 0 buy 2 at 1 one 1 time 4 ; 5 but 1 Health 1 Center 1 director 3 Lynn 1 Choafe 4 says 2 these 1 things 4- are 1 certain 4 ; 5 only 1 latex 1 condoms 4- with 2 spermicide 4 ; 4a the 1 most 1 effective 3 at 1 preventing 1 transmission 3- of 1 the 1 AIDS 1 virus 4 ; 4b will 1 be 1 available 4 ; 6 and 1 Choafe 1 says 4- students 3 will 1- be 1 required 4 to 0 undergo 1 counseling 4 the 1 first 1 time 3 they 1 come 2 to 1 the 1 clinic 4- for 1 condoms 4 ; 6 signal /u9/smg/data/radio/f2b/radio/st11/f2bst11p2 a 1 survey 2 last 1 spring 4 indicated 3 that 2 fifty 1 one 1 percent 4 of 1 Ringe 1 and 0 Latin 1 students 4 have 1 had 2 sexual 1 intercourse 4 ; 6 When 1 asked 3 what 1 contraceptives 2 were 1 used 4 the 1 last 1 time 2 they 1 had 1 sex 4 ; 5

nearly 1 half 2 said 1 condoms 4 ; 5 another 1 twenty 1 five 1 percent 4 said 1 none 4 ; 6 While 1 the 1 survey 1 showed 3 wide 1 awareness 1 of 0 AIDS 4 ; 5 many 1 students 1 say 4 it's 1 too 1 embarrassing 3 or 1 inconvenient 4 to 1 stock 1 up 2 on 1 condoms 4 at 1 the 1 corner 1 drug 1 store 4 ; 6 There 1 are 1 no 2 reported 1 cases 3 of 1 HIV 1 infection 2 at 1 the 1 school 4 ; 5 but 1 senior 2 Matt 1 Aldridge 3 is 1 worried 4 about 2 what 1 the 1 future 2 may 1 bring 4 ; 6 but 1 senior 3 Emily 1 Case 2 says 4 if 1 the 1 school 1 committee 4- tinkers 1 with 1 the 1 policy 4 ; 5 her 1 AIDS 2 peer 1 leadership 1 group 4 may 1 be 1 forced 2 to 1 resume 4 condom 2 distribution 4 ; 6 Assuming 3 the 1 Teen 1 Health 1 Center's 3 program 3 gets 1 off 2- the 1 ground 4 Case 1 says 3 she'll 1 refer 3 her 2 former 1 customers 4 to 1 the 1 clinic 4 ; 6 but 0 if 1 someone's 1 afraid 3 to 1- talk 1 to 1 adults 4 Case 1 says 4 she 1 won't 1 let 1 that 1 student 3 walk 1- away 4 empty 1 handed 4 ; 6 for 1 WBUR 4 I'm 0 Margo 1 Melnicove 4 ; 6

STORY 12

If 1 the 1 beginning 3- of 1- our 1 interview 4 was 1 any 1 indication 4 ; 5 Julie 1 Ballaga's 3- reputation 4 as 1 an 0 enthusiastic 3 accessible 4- person 4 is 1- well 1 deserved 4 ; 6 her 1 corner 1- office 1 in 1 Boston 4 has 1 a 1 fine 2 view 3- of 1 the 1 Charles 1 River 4 ; 5 but 1- a 1 poor 1 background 4 for 2 recording 4 due 1 to 0 a 1 noisy 1 heating 1 system 4 ; 6 Instead 1 of 1 asking 2 an 1 aide 3 to 0 find 1 us 1 a 1 quiet 1 office 4 ; 5 Ballaga 1 lead 1 the 1 search 4 ; 6 Fifty 1 nine 1 year 1 old 1 Ballaga 4 first 1 got 0 involved 4 in 1 environmental 1 issues 4 as 1 a 1 town 2 planning 1 and 0 zoning 1 commissioner 4 ; 5 then 1- as 1- a 1 member 3- of 0 the 1 Connecticut 1 House 1 of 1 Representatives 4 where 1 She 1 spent 3- much 0 of 0 her 2- ten 1 year 1 stint 4 on 1 the 1 Environment 3 Committee 4 ; 6 She 1 insists 2 her 1 husband's 1 work 4 as 1 an 1 oil 1 company 1 executive 4 ; 5 never 1 colored 2 her 1 thinking 4 ; 6 Ballaga 1 says 1 in 1 fact 4 She 1 taught 2- him 4 ; 5 the 1 environmental 1 way 4 ; 6 She 1 carried 1 her 1 message 3 statewide 4- in 1 nineteen 1 eighty 1 six 4 as 2 Connecticut's 2 Republican 1 candidate 1 for 1 governor 4 ; 5 against 1 incumbent 4- William 1 O'Neil 4 ; 6 O'Neil 2 trounced 1 Ballaga 4 but 1 She 1 says 1 She 1 has 1 no 1 regrets 4- about 1 having 1 run 4 ; 6 Apparently 4 her 1 politicing 2 over 1 the 1 years 4 has 1 been 1 effective 4 ; 5 it's 1 difficult 3 to 1 find 4 a 1 Ballaga 2 enemy 4 ; 6 Wanda 2 Ricurbee 4 ; 4a now 2 with 1 Connecticut's 3- environmental 1 Protection 1 Department 4 ; 5b was 1- an 1- environmental 1 lobbyist 4 when 2 Ballaga 1 was 1- a 1 legislator 4 ; 6 Head 1- of 1 governmental 1 affairs 4 for 1 the 1 Hartford 1 Chamber 1 of 1 Commerce 4 ; 4a Joseph 1 Ierna 4 ; 5b is 1 another 4- longtime 1 Ballaga 2- observer 4 ; 6 Massachusetts 1 Water 1 Resources 1 Authority 4- ; 4-a Director 3 Paul 1 Levy

4 ; 5b says 2 Ballaga 4 is 2 sincerely 1 committed 4- to 1 pro-environment 1 goals 4 ; 6 her 1 predecessor 4 ; 4a Michael 1 Deland 4 ; 5b started 2- the 1 clean-up 3- of 1 Boston 1 Harbor 4 ; 6 Levy 1 is 3- confident 4- Ballaga 4- will 1 keep 2 the 1 project 4- on 1 track 4 ; 6 Ballaga 1 has 1- already 2 passed 3 two 1 tests 4 according 1 to 1 conservationists 4 ; 6 She 1 pressed 3 for 1 Washington's 2 recent 1 rejection 4 of 0 a 1 proposed 1 dam 4 on 2 Rhode 1- Island's 2 Big 1 River 4 ; 5 and 1 insisted 1 on 1- another 1 review 4 of 1 expansion 1 plans 4 for 1 the 1 Loon 1- Mountain 2 ski 1 area 4 ; 6 Both 1 positions 4 were 1 crafted 2 by 1 Deland 4 ; 5 and 1 Ballaga 4- seems 1 determined 4 to 1 stay 3 the 1 environmental 1 course 4 ; 6 but 1- some 1 observers 4 who 1 say 2 President 2 Reagan's 4 hands 1 off 1 style 4 ; 5 allowed 1 Deland 4 plenty 1 of 1 room 4 to 1 flex 1 environmental 1 muscle 4 ; 5 wonder 2- If 0 Ballaga 3 will 1 fare 1- as 1- well 4 ; 5 under 1 the 1 hands 2- on 4- Bush 1 administration 4 ; 6 Ballaga 1 says 4 she's 1 not 1 worried 4 ; 5 and 1- denies 3- speculation 4 that 1 Chief 1 of 1 Staff 4- John 2 Sununu 4 ; 5 is 1 meddling 4- in 1 the 1 region's 3 environmental 2 affairs 4 ; 6 If 0 he 1 calls 4 Ballaga 1 says 4- she'll 1 give 1 Sununu 4- a 2- fair 1 hearing 4 ; 5 just 2 like 2 everyone 1 else 4 ; 6 for 1 WBUR 4 I'm 1- Margo 1 Melnicove 4 ; 6

STORY 13

a 1 decade 1 ago 4 there 0 were 3 three 1 reported 1 AIDS 1 cases 4- in 1- Massachusetts 4 now 2 there 1 are 3 three 1 thousand 4 ; 6 and 1 it's 1 estimated 3 that 1- every 1 day 4 six 2 more 1 Baystaters 4 become 1 infected 3 with 1- the 1 AIDS 1 virus 4 ; 6 in 1- the 1 first 1 three 1 months 4 of 1 nineteen 1 ninety 1 alone 4 ; 5 Human 1 Services 1- Secretary 4- Phillip 1 Johnston 4 says 3 two 1- hundred 3- and 1 thirty 1 one 3 men 3 women 0 and 2- children 4 ; 5 were 1- diagnosed 3- as 1 having 2 full 1 blown 1 AIDS 4 ; 6 Johnston 1 says 3 one 1 change 4 that 1- would 1 help 1 curb 2 the 1 epidemic 4 would 1- be 2 legislative 2 support 4- for 1 higher 1 taxes 4 ; 6 Without 1 additional 1 revenues 4 ; 5 Johnston 1 says 4 the 1 Public 1 Health 1 Department's 4 twenty 1 million 1 dollar 3- AIDS 1 prevention 1 programs 4 ; 5 won't 1 reach 2- enough 1 people 4 ; 6 and 1 while 3- he 1 applauded 3 President 1 Bush's 1 recent 1 remarks 4 underscoring 2- the 1 seriousness 4- of 0 the 2 AIDS 1 epidemic 4 ; 5 Johnston 1 says 4 attitudes 1 at 1 the 1 White 1 House 4 must 1 also 1 change 4 ; 6 when 1 asked 2 whether 2- any 1 changes 3 are 1 needed 3- in 1 his 1 office 4 ; 5 Johnston 1 says 4 he 1 may 1 re-evaluate 4 a 1 prohibition 3 against 2 condoms 3 in 1 state 1 prisons 4 ; 5 even 1 though 2 sex 1 among 2 inmates 4 is 2- illegal 4 ; 6 Johnston 1 says 3 given 3- prison 1 overcrowding 4 ; 5 and 1- the 1 fact 3- that 2 fifteen 1 percent 2- of 1 inmates 4 who 2 volunteered 0 to 1- be 1

tested 4 are 1 infected 3- with 1 the 1 AIDS 1 virus 4 ; 5 the 1 policy 4- deserves 3 a 1 second 1 look 4 ; 6 But 1 when 1 asked 3- if 1 the 1 state 3 should 1 reconsider 3 its 1 policy 3 against 2 giving 3 sterile 1 needles 4 to 1 I-V 2 drug 1 users 4 ; 5 Public 1 Health 1 Commissioner 4 ; 4a David 1 Mulligan 4 ; 4b says 2 no 4 ; 6 Treatment 4 ; 4a says 1 Mulligan 4 ; 4b is 1 the 1 key 4 to 1- stemming 3- the 1 spread 1- of 1 AIDS 4 among 2 addicts 4 ; 6 But 0 AIDS 2 Action 1 Committee 4- Director 3 Larry 1 Kesler 4 ; 5 says 2 more 1 aggressive 1 policies 4 and 2 much 1 more 1 money 0 are 0 needed 0 on 1 all 1 fronts 4 from 2 federal 1- agencies 4 to 1- cities 0 and 0 towns 4 ; 6 the 1 greater 1 Boston 1 area 4 ; 4a says 1 Kesler 4 ; 5b used 0 to 1 rank 2 eighth 4 among 2 major 1 cities 4 for 1 incidence 0 of 0 AIDS 4 ; 6 But 1 Kesler 1 says 3 other 1 cities 4 with 2 few 1 prevention 1 programs 4 have 1 moved 2- up 0 on 0 the 1 list 4 and 2- Boston 1 is 1 now 4- thirteenth 4 ; 6 Kesler 2 says 4 it 1 will 1 take 1 an 1 additional 4- fifteen 2- million 1 dollars 4 in 1 state 1 AIDS 2 prevention 1 funds 3 next 1 year 4 ; 5 to 1- get 1 that 1 message 3 across 4 ; 6 for 1 WBUR 4 I'm 1- Margo 1 Melnicove 4 ; 6

STORY 14

Scott 1 Harshbarger 4- is 1- no 1 stranger 3 to 0 party 1 infighting 4 He 1 challenged 3 a 1 Democratic 1 incumbent 3 in 0 an 1 unsuccessful 1 bid 1 for 1- District 1 Attorney 3 in 1 nineteen 1 seventy 1 eight 4 ; 6 Then 1 again 3 in 1 eighty 1 two 3 when 1 He 1 won 1 the 1 seat 4 ; 6 Apparently 4 the 1 former 1 Harvard 1 football 1 player 4 isn't 1- a 1 quitter 4 ; 6 but 1 He 0 and 0 James 1 Shannon 3 have 2 so 1 much 1 in 1 common 4 ; 5 that 2 many 1 observers 4- wonder 1 why 2 Harshbarger 4- wants 1 to 1 upset 3 the 1 Democratic 1 appletart 3 this 1 time 1 around 4 ; 6 Both 1 candidates 3 are 1 young 4 ; 5 Both 1 reflect 3 the 1 party's 2 liberal 1 leanings 4 and 1 support 2 abortion 1 rights 4 ; 6 but 1 Harshbarger 4 with 2 twenty 1 years 1 experience 4 in 1 criminal 3 and 0 civil 1 courts 4 ; 5 says 3 he's 1 more 1 qualified 4 to 1- be 1 the 1 state's 2 top 1 lawyer 3- than 1 Shannon 4 ; 5 who 1 was 0 a 1 Congressman 4 before 1 He 1 was 1 elected 3- Attorney 1 General 4 ; 6 as 0 an 1 example 4 the 1 two 1 term 3 District 1 Attorney 4 cites 2 what 1 He 1 calls 4 the 1 criminal 1 justice 1- system's 4 ; 5 state 2 of 2 total 1 crisis 4 ; 6 signal /u9/smg/data/radio/f2b/radio/st14/f2bst14p3 Attorney 1 General 4 James 1 Shannon 4 ; 6 Shannon 2 prides 1 himself 4 on 1 having 2- created 3- a 1 narcotics 1 unit 4 and 1- a 1 public 1 integrity 1 division 4 ; 5 which 1 are 1 waging 3 aggressive 1 wars 4 ; 4a He 1 says 4 ; 5b against 2 drugs 3 and 2- political 1 corruption 4 ; 6 Shannon's 3 best 2 known 3 ; 3a however 4 ; 4b for 1 his 1 anti-Seabrook 1 crusade 4 ; 6 that 1 battle 4- actually 1 began 3- under 1 Harshbarger 4 ; 6 who 1 filed 1 the 1 state's 2 first 1 brief 4 in 1 the 2- Seabrook 2

Power 1 Plant 2- case 4 ; 6 as 1- an 1 Assistant 1 Attorney 1 General 4- some 1 fifteen 1 years 1 ago 4 ; 6 Shannon 1 says 4- it's 1 his 1 job 4- to 1 continue 1 the 1 fight 4 ; 5 but 1 Harshbarger 3 says 4 Shannon's 1 taken 0 it 2 too 1 far 4 ; 5 because 0 he's 1 more 1 preoccupied 3 with 0 the 1 news 1 headlines 4 than 1 with 0 the 1 legal 1 merits 4 ; 6 Harshbarger 1 may 1 have 0 a 1 point 4 says 1 Pollster 2 Jerry 1 Travinsky 4 ; 5 but 0 it's 1 not 1 compelling 1 enough 4 to 0 convince 1 voters 4 to 1 unseat 3 an 1 incumbent 4 ; 6 Nor 3 ; 3a says 2 Travinsky 4 ; 4b is 1 Harshbarger's 1 claim 4 that 1 he's 1- the 1 more 1 qualified 1 candidate 4 ; 6 Travinsky 1 says 4- another 1 problem 3 for 1 Harshbarger 4 is 0 that 1 He 1 and 1 Shannon 2 are 1 cut 3 from 1 the 1 same 3- ideological 1 cloth 4 ; 5 providing 2- little 1 reason 4 for 1 people 1- to 1 switch 2 their 1 votes 4 ; 6 in 1 fact 4 it's 2- difficult 3 to 1- tell 1 the 1 candidates 3 apart 4 when 1 they 1 talk 1 about 1 the 1 issues 4 ; 6 Scott 1 Harshbarger 1 says 4 he's 1 confident 4 he'll 1 win 0 enough 2 convention 1 delegates 4 to 1 secure 1 a 1 place 3 on 1- the 1 primary 1 ballot 4 ; 6 so 1 between 2 now 3 and 1 September's 1 election 4 ; 5 Harshbarger 1 says 3 he'll 1 put 1 all 1 his 1 resources 4 into 1 distinguishing 3- himself 4 ; 5 from 2- the 1 incumbent 4 ; 4a James 1 Shannon 4 ; 6b for 2- WBUR 4 I'm 1- Margo 1 Melnicove 4 ; 6

STORY 16

Boston 1 Phoenix 1 reporter 4- ; 4a Rick 1- Kahn 4 ; 4b says 2 top 1 city 1 hall 4- and 2 police 1 officials 4 ; 5 are 1 engaged 3 in 1 a 1 deadly 1 game 2 of 1 denial 4 ; 5 giving 1 the 1 public 3 the 1 impression 4- that 1 Boston's 2 gang 1 problem 4 ; 5 is 1 confined 4 to 1 a 2 few 1 hundred 4 inner-city 1 kids 4 ; 6 But 1 Kahn 2 says 4 his 1 police 4 and 2 street 1 sources 4 ; 5 say 3 there 1 are 2 four 0 to 1 six 1 thousand 4 Boston 2 gang 1 members 4 ; 4a ages 3- nine 2 to 1 thirty 1 five 4 ; 5b who 1- are 1 competing 2 for 2- turf 4 and 2 for 2- profits 4 from 1 illicit 1 drug 1 deals 4 ; 5 across 1 the 1 city 4 and 2 beyond 4 ; 6 Kahn 2 says 2 gangs 3 use 1 the 1 latest 2 communications 1 devices 4 beepers 3 car 1 phones 3 portable 3- police 1 scanners 3- and 1 the 1 like 4 ; 5 to 1 stay 1 at 1 least 3- one 1 step 1 ahead 1 of 0 the 1 cops 4 ; 6 he 1 says 3 five 2 gangs 4 have 1 joined 2 forces 3 in 1- a 1 mega-merger 4 ; 5 giving 1 them 2 more 1 bang 4 as 1 well 1 as 2 more 1 drugs 0 and 0 guns 4 for 1 the 1 buck 4 ; 6 and 0 he 1 says 3 there 0 are 0 about 2 ten 4 up 1 and 1 coming 2 girl 2 gangs 4 ; 6 Boston 1 police 1 deputy 1 superintendent 4 William 2 Selester 4 is 1 in 1 charge 2 of 1 the 1 Mattapan 2 Roxbury 2 area 4 ; 6 Selester 1 says 2 once 3 there 1 were 3 two 1 girl 1 gangs 4 with 1 thirteen 1 members 1 each 4 ; 5 But 1- they 1 didn't 1 last 4 ; 6 he 1 says 3 mega-mergers 4 have 1 been 1 tried 4 But 2 never 1 succeeded 4 ; 5 because 1 gangs 3- can't 1 get 0 along 4 ; 6 and 0 he 1 says 3 you'd 1

have 1 to 1 count 4 ; 4a as 2- he 1 puts 1 it 4 ; 4b every 3 pain 1 in 1 the 1 neck 2- group 3 that 1 drinks 1 beer 3- on 0 a 1 street 1 corner 4 ; 5 for 1 gang 1 membership 3 to 1 approach 4 four 0 to 1- six 1 thousand 4 ; 6 Selester 1 says 4- at 1 most 4 there 1 are 1 only 3- six 1 hundred 4 real 1 gang 1 members 3- in 1- the 1 city 4 ; 5 and 0 of 1 those 3 only 2 two 1 hundred 4- are 1 hard 1 core 2 violent 1 criminals 4 ; 6 he 1 dismisses 3 claims 3- by 1 the 1 Phoenix 4 that 1 police 2 downplay 1 the 1 situation 4 ; 5 the 1 media 3 ; 3a says 1 Selester 4 ; 4b overstate 3 the 1 problem 4 ; 6 and 1- most 1 ignore 1 his 1 plea 4 not 0 to 1 publicize 4 gangs' 3- names 4 ; 6 Selester 1 says 4- such 1 publicity 3 encourages 3- gang 2 members 4 by 1 feeding 1 their 1 hunger 1 for 1 status 4 ; 6 Still 4- Selester 1 says 3- his 1 forces 4 are 1 making 1 headway 4 ; 6 Diane 1 Wilkerson 3- is 1 legal 1 counsel 4 for 1 the 1 N 1 Double 1 a 1 C 1 P's 2 Boston 1 Branch 4 and 1- an 1 inner-city 1 resident 4 ; 6 in 1 the 1 past 4 Wilkerson 1 has 1 agreed 4 with 2 Selester 2 and 0 others 4 who 1 fault 2 the 1 media 4 for 2- glorifying 1 gangs 4 ; 6 But 1- Wilkerson 3- says 4 the 1 main 1 reason 4 she 1 distrusts 1 the 1 media 4 and 0 the 1 police 4 ; 5 is 1 that 2- neither 4 have 1 been 2- telling 0 it 3- like 0 it 1 is 4 ; 6 the 1 Phoenix 1 report 4 on 2 gangs 2 Incorporated 3 ; 3a says 1 Wilkerson 4 ; 5b is 1 a 1 welcome 3 change 4 ; 6 Wilkerson 1 says 4 it's 1 no 1 wonder 3 that 1 youngsters 4 without 2 real 1 alternatives 4 ; 5 find 1- the 1 lure 2 of 2- big 1 drug 1 profits 4 hard 1 to 1 resist 4 ; 6 and 0 she 1 says 3 the 1 news 4 that 2 this 1 temptation 3 is 2 not 1 limited 4 to 0 any 1 particular 1 race 3 or 1 neighborhood 4 ; 5 also 1 comes 3 as 1 no 1 surprise 4 ; 6 for 1 WBUR 4 I'm 0 Margo 1 Melnicove 4 ; 6

STORY 17

Disbelief 4 outrage 4 guilt 4 fear 4- these 1 are 2 typical 1 reactions 4 when 1 a 1 loved 1 one 2 is 1 lost 4 to 1 random 1 violence 4 ; 6 Survivors 3 rehash 2 the 1 crime 3 over 1 and 1- over 4 ; 5 in 1 a 1 futile 1 search 3 for 1 some 1 reason 4 ; 6 poverty 4 drugs 4 easy 1 access 2 to 1 guns 4 ; 5- or 1 perhaps 2 this 1 explanation 4- of 1 gang 1 mentality 4 from 1 Josephine 1 Johnson 4 ; 6 Johnson's 3 fifteen 1 year 1 old 1 son 2 Charles 4 was 1 stabbed 2 in 1 the 1 heart 4 in 1 a 1 Roxbury 1 alley 4 by 1- a 1 teen-age 1 acquaintance 4 a 1 year 1 and 0 a 1 half 1 ago 4 ; 6 she 1 says 3 her 1 anger 4 and 2 her 1 memories 4 have 1 kept 1 her 1 alive 4 ; 6 Johnson 3- tends 1 a 1 small 1 table 3 in 1 her 1 apartment 4 ; 5 which 1 she 1 lovingly 1 calls 3 Charles' 2 stand 4 ; 6 on 1 it 4- she's 1 carefully 1 arranged 3 ribbons 4- from 1 his 1 funeral 1 wreaths 4 ; 5 and 1 many 2 of 1 her 1 son's 4- favorite 1 things 4 ; 6 Johnson 2 won't 1 let 1 the 1 outside 1 world 3- forget 1 Charles 2 either 4 ; 6 she's 1 a 1 dedicated 1 member 3 of 2 parents 2 of 1

Murdered 1 Children 4 ; 5 a 1 group 3- whose 2 growing 1 ranks 4 are 1 pushing 1 hard 4 for 1 juveniles 4 who 1 commit 2 violent 1 crimes 4 to 1 be 1 tried 2 as 1 adults 4 ; 6 Twenty 1 seven 1 year 1 old 3 Michael 1 Farrow 4 says 1- at 1 first 3 He 1 considered 4 taking 1 the 1 law 3 into 1 his 1 own 1 hands 4 to 0 get 0 at 1 the 1 people 3 who 1 ruined 2 his 1 life 4 ; 6 the 1 former 1 construction 1 worker 4 unwittingly 1 walked 2 into 1 a 1 gang 1 shoot-out 4 near 1 a 1 Roxbury 2 housing 1 project 4 last 2 August 4 ; 6 He 1 lost 2 an 1 eye 4 ; 6 He 1 also 1 lost 2 his 1 job 4 apartment 3- and 1 girlfriend 4 ; 6 but 1 Farrow 1 says 3 now 2 He 1 realizes 4 two 1 wrongs 4- don't 1 make 1- a 1 right 4 ; 6 and 0 He 1 says 3 perhaps 1 his 1 pain 4 spared 3 someone 1 else 4 ; 6 of 1 everyone 3 interviewed 3 for 1- this 1 report 4 Alice 1 Moore's 2 had 1 the 1 most 1 time 4 to 1- learn 2 to 1- cope 1 with 1 her 1 loss 4 ; 6 now 2 she's 1- able 4- to 1 recall 2 her 1 daughter 4 without 2 breaking 1 down 4 ; 6 are 1 those 2 pictures 2- of 1 her 4 on 0 the 0 organ 4 ; 6 Moore's 1 helping 2 a 1 support 1 group 4 get 0 off 2 the 1 ground 4 for 2- parents 1 like 1- herself 4 ; 6 It's 2 too 1- bad 3 there 1 are 1 so 1 many 1 members 4 ; 4a she 1 says 4 ; 5b but 0 It's 2- comforting 3 to 1- know 4 you 1 are 1 not 1 alone 4 ; 6 for 2- WBUR 4 I'm 1- Margo 1 Melnicove 4 ; 6

STORY 18

Jonie 1 says 3 the 1 old 2 ways 4 are 1 also 1 important 4 ; 5 Living 1 off 1 the 1 land 4 special 1 ceremonies 4 respecting 1 one's 1 elders 4 and 1 being 1 part 3 of 1 an 1- enormous 1 extended 1 family 4 ; 6 So 1 last 1 winter 4 after 2 going 1 home 1 for 1 Christmas 4 ; 5 she 1 re-enrolled 3 In 0 a 1 public 1 school 4 near 1 her 1 reservation 4 ; 6 but 1 after 3- three 1 days 2 of 1 classes 4 she 1 says 3 she 1 realized 3 the 1 importance 3 of 1 getting 3- the 1 best 1 Education 1 possible 4 ; 5 and 1 returned 4- to 1 Northfield 1 Mount 1 Herman 4 ; 6 it's 1 nothing 1 new 4 for 1 reservation 1 children 4 who 1 attend 2 boarding 1 schools 4 to 1 feel 1 torn 4 ; 5 but 1 there's 1 no 1 comparison 4 between 1 today's 1 prep 1 schools 4 says 2 Northfield 1 Mount 1 Herman's 1 advisor 4 for 1 native 1 American 1 students 4 Tom 1 Sorcee 4 ; 5 and 1 the 1 old 3 missionary 4- and 2 government 1 run 2 boarding 1 schools 4 ; 5 where 1 children 4- were 1 beaten 4 for 1 speaking 1 native 1 languages 4 ; 5 and 1 the 1 goal 3 was 2 to 1 Christianize 4 or 1 Americanize 1 them 4 ; 6 Sorcee 1 says 3 his 1 goal 4 is 1 to 1 provide 3 a 2 home 1 away 1 from 1 home 4 ; 5 and 1 apparently 4- he's 1 succeeding 4 ; 6 Sorcee 1 says 4- during 1 the 2 five 1 years 4 he 1 and 1 his 1 wife 4 who 1 is 1 Navajo 4 have 1 been 1- at 1 Northfield 1 Mount 1 Herman 4 ; 5 the 1 student 1 body's 1 included 3- about 1 a 1 dozen 2 native 1 Americans 2 each 1 year 4 ; 5 and 1 only 1 one 4 has 1 dropped 1 out 4 ; 6 the 1 network 2 is 1

growing 4 as 2 more 1- and 1- more 1 prep 1 schools 3 recruit 2 native 1 American 1 students 4 ; 6 Doctor 1 Joseph 1 Curry 4 headmaster 1 at 2 Cushing 1 Academy 4 In 1 Ashburnham 1 Massachusetts 4 ; 5 is 1 greatly 1 responsible 3 for 1- the 1 trend 4 ; 6 Curry's 2 determined 4 to 1 build 1 a 1 tuition-free 1 prep 1 school 3- In 1 the 1 southwest 4 ; 5 just 1 for 1 native 1 Americans 4 ; 6 Meanwhile 4 Cushing 1 Academy 4 sponsors 2 a 1 six 1 week 3 junior 1 high 2 summer 1 school 1 there 4 ; 5 based 2 on 1 the 1 prep 1 school 1 model 4- of 1 rigorous 1 academics 4 all 1 day 4 ; 5 study 1 hall 3- at 1 night 4 ; 5 sports 1 and 1 field 1 trips 3 on 1 weekends 4 ; 6 Enrollment 2 has 1 grown 4 ; 5 from 2 fifty 1 students 3 In 2 nineteen 1 eighty 1 eight 4 to 1 more 1 than 2 three 1 hundred 3 this 1 summer 4 ; 6 Curry 1 says 3 many 1 native 1 American 2 families 4 thrilled 3- to 1 see 1 their 1 children 3 turned 1 on 2 by 1 school 2 for 1- a 1 change 4 ; 5 now 1 want 1 to 1 enroll 1 them 3 In 1 prep 1 school 3 year-round 4 ; 6 In 2 fact 4 says 2 Doctor 1 Betty 1 Haskins 4 it's 1 essential 4 for 1 native 1 American 1 children 4 to 1 learn 1 how 1 to 1 walk 2 proudly 4- In 1 both 1 worlds 4 ; 6 and 1 the 1 director 3 of 1 the 1 American 1 Indian 1 Program 3 at 1 Harvard's 2 Graduate 1 school 1 of 1 Education 4 applauds 2 anything 3 prep 1 schools 3 can 1 do 4- to 1 help 3 pave 1 the 1 way 4 ; 6 the 1 Cherokee 1 educator 2 says 3 each 1 child's 2 future 4 as 1 well 1 as 2 fragile 1 treaty 1 rights 4 ; 5 are 1 at 1 stake 4 ; 6

STORY 19

Lincoln's 1 reputation 3 as 1 a 1 town 3 with 1 a 1 social 1 conscience 4 arises 2 primarily 2 from 1 two 1 traditions 4 ; 5 a 1 commitment 3 to 1 land 1 conservation 3 That 2 predates 3 the 1 environmental 1 movement 4 by 2 several 1 decades 4 ; 5 and 1 a 1 commitment 3 to 1 education 4 ; 6 Lincoln 1 voters 4 who 1 tend 1 to 1 be 2 highly 1 educated 1 themselves 4 ; 5 have 1 approved 3 a 1 series 4 of 2 Proposition 1 Two-and 1 overrides 4 ; 5 and 1 spend 2 more 1 per 1 pupil 4 than 2 most 2 Massachusetts 1 communities 4 ; 6 the 1 town 2 has 1 shared 2 its 1 riches 4 with 2 black 1 youngsters 1 from 1 Boston 4 since 2 METCO's 1 inception 3 twenty 1 five 1 years 1 ago 4 ; 6 and 1 each 1 year 2 since 1 nineteen 1 seventy 1 five 4 Lincoln's 2 METCO 1 enrollment 3 has 1 comprised 2 twenty 1 percent 4 of 1 the 1 student 1 body 4 ; 6 but 2 no 1 town's 1 an 1 island 4 ; 5 not 1 even 1 Lincoln 4 ; 5 says 2 school 1 committee 1 chairman 4 Henry 1 Morgan 4 ; 6 Morgan 2 stopped 2 short 4 of 1 saying 2 racism 4 may 1 underlie 2 recent 1 complaints 3 about 1 METCO 4 ; 5 preferring 2 instead 2 to 1 say 3 That 2 some 1 people 3 may 1 be 2 insensitive 4 to 1 the 1 implications 3 of 1 race 4 ; 6 Morgan's 2 not 1 the 1 only 1 Lincolnite 4 who's 2 choosing 1 words 2 carefully 4 ; 6 Irene 1 Rice 1

says 4 it's 1- as 1- if 1 METCO's 4- a 1 sacred 1 cow 4 ; 5 which 2 few 1 Lincoln 2 parents 4- have 1 dared 3 to 1 question 1 publicly 4 ; 6 One 1 of 1 Rice's 1 complaints 4 is 2 That 2 teachers 1 failed 1 to 1 recognize 4- her 1 son's 2 learning 1 disability 4 ; 6 Apparently 4- because 2 they 1 were 2 preoccupied 3 with 1 other 1 children 4 ; 6 Rice 1 points 1 out 4 That 1 while 2 METCO 1 youngsters 3 comprise 3 one-fifth 1 of 1 the 1 students 4 ; 5 they 1 require 3 nearly 2 half 4- of 1 the 2 individual 2 education 1 plants 4 ; 6 other 1 Lincoln 1 parents 3- worry 3 That 1 the 1 schools 3 are 1 losing 4- their 2 academic 1 edge 4 ; 6 Tracking 1 students 3 according 1 to 1 ability 4- is 2 frowned 1 upon 4 by 2 Lincoln 1 teachers 2 and 1 school 1 administrators 4 ; 5 because 1- it's 2 been 1 shown 4 to 1 discriminate 4- against 2 minorities 4 ; 6 but 1 school 1 committee 1 member 4 Jennifer 1 Donaldson 4 ; 5 wonders 2 if 1 the 1 system's 2 anti-tracking 1 philosophy 4 ; 5 shortchanges 1 some 1 students 4 for 1 the 1 sake 3 of 1 their 1 METCO 2 classmates 4 ; 6 but 1 Donaldson 1 says 4 what 1 may 1 have 1 forced 2 most 1 METCO 1 critics 2 out 1 of 1 the 1 closet 4 is 1 the 1 statewide 2 economics 1 squeeze 4 ; 6 the 1 state 3 reimburses 1 Lincoln 4 for 2 METCO 1 related 1 costs 4 ; 5 but 1 those 1 funds 3 face 1 a 1 five 1 percent 1 cut 4 ; 6 Meanwhile 4 Lincoln 1 enrollment 3 is 1 rising 4 ; 5 requiring 2 two 1 more 1 classrooms 3 next 1 year 4 ; 6 That 1 expense 3 and 2 others 4 could 1 be 1 avoided 4 some 2 taxpayers 3 say 4 ; 5 by 1 downsizing 1 METCO 4 ; 6 Estimates 4 vary 1 widely 4 on 2 how 1 much 1 money 4 could 1 be 1 saved 4 ; 6 but 1 no 1 matter 1 what 2 the 1 cost 4 school 1 committee 1 member 2 Mickey 1 Lipsey 1 says 4 you 1 can't 3 put 1 a 1 price 3 on 1 the 1 value 4 of 1 well- 2 integrated 1 classrooms 4 ; 6

STORY 20

Carol 2 Stuart's 1 brother 4 Carl 1 DiMaiti 4 told 1 reporters 3 and 1- a 1 crowd 2 of 1 onlookers 4 ; 5 it's 1 only 1 fitting 4 that 1 we're 1 at 2 Brigham 1 and 1 Women's 1 Hospital 4 in 1 Mission 1 Hill 4 ; 6 That's 1 where 2 Carol 1 Stuart 3 and 1- her 1 husband 3 attended 1 a 1 childbirth 1 class 1 last 1 Fall 4 ; 5 shortly 1 before 2 Charles 1 Stuart 3 apparently 2 orchestrated 2 his 1 wife's 1 murder 4 outside 2 a 1 nearby 1 housing 1 project 4 ; 6 He 1 then 2 blamed 1 the 1 crime 3 on 1 a 1 black 1 man 4 ; 6 in 1- an 1 effort 2 to 1 ease 3 the 1 racial 1 tensions 3 that 1 erupted 4 as 1 well 1 as 2 their 1 own 2 pain 4 Carol 1 Stuart's 1 family 3 launched 1 a 1 foundation 4 to 1 provide 2 college 1 scholarships 4 For 1 Mission 1 Hill 1 residents 4 ; 5 and 1 to 1 support 3 programs 4 that 1 combat 2 racism 4 ; 6 Stuart's 1 mother 4 Evelyn 1 DiMaiti 4 ; 5 said 2 she's 1 overwhelmed 4- by 1 the 1 talented 1 people 4 who 1 volunteered 3 For 1 the 1 foundation's 1 board 4 ; 6 But 1 first 4 the 1 trustees 3 will 1 review 2 scholarship 1 candidates

4 ; 6 More 1 than 1 one 1 hundred 1 applications 3 have 1 been 1 distributed 3 in 1 just 1 two 1 weeks 4 ; 6 DiMaiti 1 says 3 awards 3 will 1 be 1 made 2 in 1 August 4 ; 6 the 1 trustees 2 also 1 have 1 some 2 fundraising 2 to 1 do 4 ; 5 About 2 ten 1 thousand 1 individuals 4 and 2 several 1 businesses 4 have 1 made 1 contributions 4 ; 5 with 1 the 1 largest 3 thus 1 far 4- from 2 Brigham 1 and 1 Women's 1 Hospital 4 ; 6 the 1 fund 2 now 1 totals 3 nearly 2 five 1 hundred 1 thousand 1 dollars 4 ; 5 But 1 Carl 1 DiMaiti 1 says 3 another 2 half 1 a 1 million 3- is 1 needed 4 For 1 the 1 Carol 1 DiMaiti 1 Stuart 1 foundation 4 to 1 become 2 a 1 lasting 1 legacy 4 ; 6 For 1 WBUR 4 I'm 1- Margo 1 Melnicove 4 ; 6

STORY 21

in 2 three 1 years 4 Act 1 Up 2 has 1 grown 4 from 1 one 1 group 3 in 1 New 1 York 1 City 4 ; 5 to 1 more 1 than 2 sixty 1 chapters 4 across 2 the 1 U 3 and 1 abroad 4 ; 5 whose 2 trademark 1 slogan 4 is 2 silence 4 equals 1 death 4 ; 6 the 1 movement 2 is 2 credited 4- with 2 forcing 1 government 1 in 1 business 4 ; 5 to 1 allow 3- AIDS 1 patients' 4 access 3 to 1 experimental 1 drugs 4 ; 6 to 1 bring 1 down 2 the 1 cost 3 of 2 A 4 the 1 only 1 approved 2 AIDS 1 drug 4 ; 5 and 1 to 1 expand 3 Insurance 1 coverage 4 for 2 AIDS 1 related 1 treatment 4 ; 6 Alan 1 Fine 3 heads 2 the 2 Harvard 1 AIDS 1 Institute 4 ; 6 He 1 says 1- it's 1 a 1 pity 4 that 2 Act 1 Up's 1 protest 4 at 1 the 1 closing 1 session 3 of 1 the 1 International 2 AIDS 1 Conference 4 ; 5 which 1 drowned 1 out 2 Health 1 and 1 Human 1 Services' 1 secretary 3 Louis 1 Sullivan 4 ; 5 also 1 seem 1 to 1 drown 1 out 4 the 1 tone 3 of 1 the 1 rest 3 of 1 the 1 Conference 4 ; 6 Still 3 Volts 2 admits 4 that 1 sometimes 4- Act 1 Up 3 makes 1 him 1 cringe 4 ; 6 Most 1 notably 1 when 1 a 1 protester 3- at 1 New 1 York's 2 Saint 1 Patrick's 1 Cathedral 4 spit 1- out 2 a 1 communion 1 wafer 4 stomping 1 it 1 into 1 the 1 floor 4 ; 6 the 1 AIDS 1 Action 1 Committee 3 remain 1 neutral 4 during 1 a 1 more 1 recent 1 protest 4 outside 1 Boston's 2 Holy 1 Cross 1 Cathedral 4 ; 5 organized 3 by 1 Act 1 Up 3 and 2 two 1 other 1 groups 4 ; 5 who 1 say 3- the 1 church 3 has 2 no 1 right 3 to 1 try 1 to 1 influence 3 public 1 policy 4 on 2 safe 1 sex 1 education 4- and 2 condoms 1 distribution 4 ; 6 As 1 eleven 1 priests 3 being 2 ordained 3 in 1 the 1 Cathedral 4 ; 5 prostrated 1 themselves 4 before 2 Bernard 1 Cardinal 1 Law 4 ; 5 protesters 1 outside 4 staged 2 a 1 so-called 1 die-in 4 ; 5 by 2 lying 1 in 1 the 1 street 4 ; 6 Chants 1 of 2 shame 4 shame 4 greeted 1 worshippers 4 leaving 1 the 1 church 4 ; 5 As 1 did 2 condoms 4 thrown 1 their 1 way 4 ; 6 Leonard 1 Zacom 3 of 1 the 2 Anti-Defamation 1 League 4 ; 5 who 1 generally 1 applauds 4 Act 1 Up's 1 causes 4 ; 5 says 2 shame 1 on 1 the 1 protesters 4 for 1 their 1 tactics 4 ; 6 Act 1 Up 1 Boston's 3 targeted 2 not 1 only 2 Cardinal

1 Law 4- but 2 Massachusetts 1 General 1 Hospital 4 Harvard 1 Medical 1 School 4 and 1 the
1 John 1 Hancock 2 Insurance 1 Company 3 in 1 the 1 past 4 ; 6 Its 2 next 1 demonstration 4
says 2 one 1 organizer 4 ; 5 is 1 likely 2 to 1 be 2 at 1 the 1 state 1 house 4 ; 6 for 1 WBUR 4
I'm 1- Margo 1 Melnicove 4 ; 6

STORY 23

It 1- has 1- not 2 been 1 disputed 4 during 2 the 1 two 1- week 1 long 1 trial 4 that 2 last 1 April
4 Lani 1 Lee 1 Gilchrist 4 fired 2 several 1 shots 4 at 1- his 1 boss 4 ; 5 George 1 Cook 4 ; 5
and 1 pistol-whipped 3 the 1 executive 3 as 1 he 1 begged 2 for 1 help 4 ; 6 But 1 Gilchrist's 1
lawyers 3 maintain 4 he 1 is 2 not 1 guilty 4 by 2 reason 3- of 2 temporary 1 insanity 4 ; 6 the
1 shooting 1 occurred 3 the 1 day 1 after 1 Gilchrist 1 was 1 fired 4 from 1 his 2 stockbroker 1
job 4 ; 5 for 1 poor 1 sales 4 ; 6 the 1 defense 3 argued 3 that 1 Gilchrist 4 who 1 is 1 black 4
suffers 1 from 1 ongoing 2 racial 1 paranoia 4 ; 5 and 1 that 1 losing 3 the 1 job 3 of 1- his 1
dreams 4 sent 1 Gilchrist 3- into 1 an 2 uncontrollable 1 frenzy 4 ; 6 But 1- the 1 prosecution
3 maintains 4 the 1 murder 3 was 2 premeditated 4 ; 5 committed 3 in 1 the 1 coldest 2 of 1
blood 4 ; 6 If 1 Gilchrist 3 had 1 truly 2 had 1- a 2 psychotic 1 episode 4 the 1 prosecution 1
argued 3 he 1 would 2 not 1 have 1 waited 1 a 1 day 4 before 2 killing 3 his 1 boss 4 ; 6 for 1
WBUR 4 I'm 1- Margo 1 Melnicove 4 ; 5 at 2 Suffolk 1 Superior 1 Court 4 ; 6 STORY 24

On 2 one 1 side 2 of 1 the 1- debate 4 before 1 the 1 Health 1 Care 1 Committee 4- are 1
those 3 who 1 claim 4- that 2 resident 1 physicians 4 need 2 legal 1 protection 4 ; 6 Residents
2 routinely 1 work 4- ninety 3 to 1 one 1 hundred 1- and 1- ten 3 hours 2 per 1 week 4 ; 5
Cambridge 2 Hospital 1 Intern 4- David 1 Ray 4 ; 5 applauds 3 the 1 proposal 4- to 1 cap 2
the 1 work 1 week 4 at 2 seventy 2 two 1 hours 4 ; 6 On 1 the 1 other 1 side 1 of 1- the 1-
controversy 4- are 2 training 1 directors 4 from 2 several 1 hospitals 4 ; 5 who 1 testify 3- that
1 they 1 are 1 voluntarily 4- improving 3 residents' 1 working 1 conditions 4 ; 5 and 1 don't 2
need 3 a 2 legislative 2 mandate 4 ; 6 Sponsor 1 of 1- the 1 bill 4 ; 4a Boston 1 Representative
2 John 1 McDonnough 4 ; 5b says 3- he 1- does 1 not 1 expect 2 the 1 measure 3 to 1 win 1
approval 4- this 1 session 4 ; 5- But 2 McDonnough 1 says 2 the 1 bill 4 puts 2 hospitals 1 On
1 notice 4 ; 5 to 1- take 2 better 2 Care 4 of 2 their 1 doctors 4 in 1 training 4 ; 6 For 1 WBUR
4 I'm 1- Margo 1 Melnicove 4 ; 6 STORY 25

Forty 1 two 1 year 1 old 2 Lani 1 Gilchrist 4 is 1 pleading 1 insanity 4 ; 5 in 1 the 1 shooting 1 death 3- of 2 George 1 Cook 4 last 2 April 4 ; 6 Cook 1 was 1 killed 3 in 2 Merrill 1 Lynch's 2 downtown 1 Boston 1 office 4 ; 5 the 1 day 1 after 2 he 1 fired 1 Gilchrist 4 For 1 poor 1 sales 1 performance 4 ; 6 Cambridge 1 Psychologist 3 Doctor 1 Guy 1 Seymour 4 ; 5 testified 3- that 2 Gilchrist 4- is 1- in 2 constant 1 turmoil 4 that 1 he 1 believes 3 people 1 are 1 out 2 to 2 get 1- him 4 because 2 he 1 is 1 black 4 ; 5 that 1- it 1 doesn't 2 take 1 much 2 to 1 push 2 Gilchrist 4 over 1 the 1 edge 4 ; 5 and 1 that's 2 what 1 happened 4 said 2 Seymour 4 the 1 day 2 of 1 the 1 shooting 4 ; 6 Seymour 1 said 3- severe 1 stress 4 caused 2 Gilchrist 4- to 1 lose 1 touch 2 with 1 reality 4 ; 6 Later 2 this 1 week 4 the 1 district 1 attorney 3 is 1 expected 2 to 1- call 2 on 1 experts 4 ; 5 to 1 testify 3 that 2 Gilchrist 4- knew 2 exactly 2 what 1 he 1 was 1 doing 4 when 1 he 1 allegedly 3 killed 1 Cook 4 ; 6 For 2 WBUR 4 I'm 1- Margo 1 Melnicove 4 ; 6

STORY 26

Lani 1 Lee 1 Gilchrist 4 is 1 charged 3 with 2 first 1 degree 1 murder 3- in 1 the 1 shooting 2 of 1- a 1 Merrill 1 Lynch 2 Vice 1 President 4 ; 4a George 1 Cook 4 ; 4b in 1 the 1 executive's 2 downtown 1 Boston 1 office 4 ; 6 Cook 1 was 1 killed 3 last 2 April 4 the 1 day 1 after 1 he 1 fired 2 Gilchrist 4 reportedly 3 for 1 poor 1 sales 1 performance 4 ; 6 Today 4- eighteen 1 people 3- asked 2 Judge 1 John 1 Owen 4 to 1 excuse 1 them 4 from 1 the 1 jury 1 pool 4 ; 5 Most 1 of 1- the 1 requests 3 were 1 granted 4 ; 6 the 1 trial 3 is 1 expected 1 to 1- last 3 three 1 to 1- five 1 weeks 4 and 1 is 2 bound 1 to 1 be 2 dramatic 4 ; 6 the 1 prosecuting 1 attorneys 0 has 0 said 1 of 1 Gilchrist 4 they 1- don't 1- come 0 any 0 more 3 violent 1 and 2 vicious 3 than 1 this 1 one 4 ; 6 But 2 Gilchrist's 1- family 1 and 1 friends 4 described 2 the 1 forty 1 two 1- year 1- old 1 bachelor 4 as 2 meek 4 and 1 mild 4 ; 6 Gilchrist 2 has 1 pleaded 1 not 1 guilty 4 ; 5 his 1 attorneys 3 will 1 not 2 disclose 3 their 1 defense 4 ; 5 But 1 they 2 are 1 likely 3 to 1 argue 3 that 1 Gilchrist 4 who 1 is 1 black 4 ; 5 was 2 unable 1 to 1 control 1 himself 4 ; 5 in 1 part 3 because 1 he 1 viewed 2 his 1 firing 4 as 2 racially 1 motivated 4 ; 6 for 2- WBUR 4 I'm 1- Margo 1 Melnicove 4 ; 6

STORY 27

Auto 1 insurance 4- was 2 overhauled 3 last 1 year 4 ; 6 Now 1 motorists 3 are 2- paid 1 directly 4 for 1 repair 1 costs 4 ; 5 based 2 on 1 an 1 insurer's 2 appraisal 4 ; 6 Consumers 1- are 1- referred 3 to 1- several 1 Body 1 shops 4 that 1- will 1 do 1 the 1 work 4 for 1 that 1 amount 4 ;

5 but 1- are 1- free 2 to 1- shop 1 around 4 ; 6 the 1 Massachusetts 3 Auto 1 Body 1 Association 4 ; 5 says 2 the 1 law 3 promotes 2 price-fixing 4 between 1 insurance 1 companies 4- and 2 so-called 1 referral 1 shops 4 ; 5 and 1- is 1 putting 3 independents 4 out 1- of 1 business 4 ; 6 but 2 consumer 1 attorney 3- for 1- the 1 Massachusetts 2 Public 1 Interest 1 Research 1 Group 4 Josh 2 Cradka 4 ; 5 says 2 independents 4 are 2 trying 3 to 2- chip 1 away 4 at 2- Auto 1 insurance 3 reform 4 ; 6 the 1 Auto 1 Body 1 Association 3 wants 2 to 1- limit 3- the 1 kinds 3 of 1 supplemental 1 payments 4 a 1 referral 1 shop 3 may 1 receive 4 ; 5 Cradka 1 says 4 that 2 will 1 discourage 2 any 1 shop 4 from 2 participating 3 in 1 the 1 referral 2 system 4 ; 6 for 2 WBUR 4 I'm 1- Margo 1 Melnicove 4 ; 6

STORY 28

the 1 city's 1 contract 4 with 1 some 2 two 1 hundred 1 and 1 seventy 3- B 1 resident 1 physicians 4 ; 5 expires 3 June 1 thirtieth 4 ; 6 It 1 guarantees 2 support 1 staff 4 ; 6 a 1 new 1 proposal 4- does 1 not 4 ; 6 Doctor 1 Maria 1 Patero 3 heads 2 the 1 residence 2 union 4 ; 6 She 1 says 3 physicians 2 in 1 training 4- work 2 ninety 3 to 1- one 1 hundred 1- and 0 twenty 2 hours 1 per 1 week 4 ; 5 and 1 won't 1 settle 2 for 0 a 1- contract 4 that 1 does 2 not 1 promise 4 ancillary 1 services 4 ; 6 Patero 1 says 3 the 1 city 3 must 2 find 1 the 1 money 4- to 1- provide 2 residents 4 with 2 basic 1 support 4 ; 5 but 1 nearly 2 eighty 4 B 4- Mattapan 3- and 1 Long 1 Island 2 hospital 2 workers 4 received 2 pink 1 slips 2 last 1 week 4 ; 5 and 2 city 1 officials 1 say 4- next 1 year's 4 bare 1 bones 1 budget 4 ; 5 will 1 necessitate 4- more 1 layoffs 4 ; 6 for 2- WBUR 4 I'm 1- Margo 1 Melnicove 4 ; 6

STORY 29

According 2 to 1 UNICEF 4 the 1 World 1 Health 1 organization 4 and 1 other 2 international 1 monitoring 1 groups 4 ; 5 one 1- of 1 every 2 three 1 people 3 who 1 die 1 each 1 day 4 ; 5 is 1- a 1 child 4 under 1 the 1 age 2 of 1- five 4 ; 6 each 1 and 1 every 1 day 4 preventable 1 diseases 4- such 1 as 2- tetanus 3- and 1 measles 4 ; 5 kill 1 almost 2 eight 1 thousand 1 children 4 ; 5 Dehydration 4 kills 3 nearly 2 seven 1 thousand 4 ; 6 pneumonia 4- more 1 than 2- six 1 thousand 4 ; 6 At 1 present 1 rates 4 almost 2 one 1 hundred 0 and 0 fifty 2 million 1 children 3 will 1 die 4 during 2 the 1 nineteen 1 nineties 4 ; 6 the 1 statistics 4 in 1 short 4 are 1 staggering 4 ; 6 alex 2 palacios 4 is 2 unicef's 2 washington 1 lobbyist 4 ; 6 he 1- and 1 other 3 child 1 advocates 3- yesterday 4 ; 5 called 1 on 2 president 2 bush 3 to 1- sign 3 the 1

convention 1 on 1 the 1 rights 2 of 1 the 1 child 4 ; 6 the 1 treaty's 1 on 1 the 1 agenda 4 for 1 this 1 weekend's 3 world 1 summit 4 for 1 children 4 at 1 the 1 U 4 ; 6 palacios 1 says 4 he's 1 received 2 informal 1 word 3 from 1- the 1 white 1 house 4 ; 5 that 1- the 1 president 3 will 1 sign 2 the 1 document 2 sunday 4 ; 5 but 1- it's 1- unclear 4- if 2 bush 3 will 2 push 1 hard 4 for 2- senate 3- ratification 4 ; 6 calling 2 the 1 treaty 3 the 1 magna 1 carta 3- for 1- children 4 ; 5 palacios 1 says 4 speedy 1 approval 2 is 1 critical 4 ; 6 he 1 admits 4 the 1 treaty 1 alone 4 won't 1 change 2 children's 2 lives 4 ; 5 but 1- says 3 it 2 can 1 change 3 attitudes 1 and 1- policies 4 ; 5 about 2 children's 2 needs 3- and 1- rights 4 ; 6 Palacios 1 says 4- the 1 treaty 1 also 3 can 1- add 2 muscle 3 to 1- the 1 argument 4 ; 5 that 1- governments 4 must 1 re-arrange 3 priorities 1 and 1 resources 4 to 1 put 2- children 2 first 4 not 1 last 4- on 1 the 1 list 4 ; 6 Tufts 1 University 1 President 4 Jean 1 Meyere 4 ; 5 makes 2 that 1 point 3 with 1 his 1 own 1 statistics 4 ; 6 Meyere 1 says 4 every 1 chief 1 of 1 state 4 will 1 profess 4 his 1- or 1- her 1 love 4- for 1- children 4 ; 5 but 1 their 1 budgets 2 primarily 1 reflect 4 a 1 love 1- affair 3 with 1 the 1 arms 1 race 4 ; 6 According 1 to 1 UNICEF 4 twenty 1 thousand 4 of 1 the 2 forty 1 thousand 1 children 4- who 1 die 2 each 1 day 4 ; 5 could 1 be 1 saved 4 by 1 expanding 3 immunization 3- and 1- nutrition 1 programs 4 ; 5 At 1- a 1- cost 3 of 2 two 1 and 1- a 1 half 2 billion 1 dollars 4- a 1 year 4 ; 6 the 1 same 1 amount 4 says 1 UNICEF 4 that 1 the 1 worlds 1 nations 4- spend 1 on 1 weapons 4- each 1 day 4 ; 6 for 2- WBUR 4 I'm 1- Margo 1 Melnicove 4 ; 6 STORY 30 in 1 nineteen 1 seventy 1 four 4 Jim 2 Grilsh 4 was 1- a 2 healthy 1 father 1 of

1 two 3 working 1 for 1- Boston 1 Edison 4 ; 5 when 2 suddenly 4 at 1 the 1 age 2 of 1 thirty 1 seven 4 He 1 lost 2 his 1 hearing 4 overnight 4 ; 6 to 1- this 1 day 4 doctors 3 don't 2 know 1 why 4 ; 6 Like 1 other 2 late 1 deafened 1 adults 4 ; 5 Grilsh's 1 loss 4 isolated 1 him 4 ; 6 He 1 doesn't 1 identify 4 with 1 people 2 who 0 are 2 born 1 deaf 4 ; 5 who 1 have 2 created 3 their 1 own 3 community 1- and 0 language 4 ; 6 Grilsh 2 says 4 he's 1- a 1 product 3 of 1 the 1 hearing 1 world 4 ; 5 and 1- it's 1 frustrating 4 to 1- no 1 longer 2 be 1- able 2 to 1 participate 3 fully 4 ; 6 Grilsh 2 hasn't 1 learned 3 sign 1 language 4 because 2 everyone 1 He 1 knows 4- can 1 hear 4 ; 6 He 1 tries 3 to 1 read 1 lips 4 ; 5 which 1- is 2 difficult 3 at 1 best 4 ; 5 and 1 has 1 his 1 family 3 and 0 friends 3 communicate 1 with 1 him 4 in 1 writing 4 ; 6 STORY 31

Police 1 say 3 Elizabeth 1 Levy 4 was 1 driving 4- fifty 1 miles 1- an 1 hour 4 in 1- a 1 thirty 1 mile 2 zone 4- around 1 midnight 4 ; 5 on 1 July 1 fifteenth 4 ; 5 when 1 she 1 lost 2 control 3- of 1 her 1 car 4 ; 6 she 1- hit 0 a 1 parked 2 auto 4 ; 5 then 1- a 1 utility 1 pole 4 ; 6 Police 1

also 1 say 4- Levy's 2 blood 1 alcohol 1 level 4- was 2 twice 2 the 1 legal 1 limit 4 ; 6 It 1- was 1 her 1 second 3 drunk 1 driving 2 arrest 4 ; 6 Levy 1 was 1 eight 1 and 1- a 1- half 1 months 1 pregnant 4 at 2 the 1- time 2 of 1 the 1 accident 4 ; 5 the 1 next 1 day 4- her 1 baby 2 was 1 stillborn 4 ; 6 Middlesex 1 County 2 District 1 attorney 2 Scott 1 Harshbarger 4 ; 5 says 3 state 2 law 4 and 1- the 2 facts 2 of 1 the 1 case 4 give 1- him 2 no 1 choice 4 but 1 to 1- prosecute 2 Levy 4 for 2 motor 1 vehicle 1 homicide 4 ; 6 but 2 some 1 critics 1 say 4 Harshbarger 4 may 1 be 2 throwing 1 the 1 book 2 at 1 Levy 4 for 2 political 1 reasons 4 ; 6 They 1 suggest 4- that 1 Harshbarger 4 ; 4a who's 1 running 3- for 2 state 1 attorney 1 general 4- in 1 nineteen 1 ninety 4 ; 5b and 1 who's 1 known 3 as 0 an 1- abortion 1 rights 1 advocate 4 ; 5 may 1- be 1 pursuing 1 the 1 case 4 to 1- curry 2 the 1 favor 3 of 1- voters 4 who 1 support 2 fetal 1 rights 4 ; 6 but 1 Harshbarger 4- says 4 the 1 case 3 has 2 nothing 1 to 1 do 2 with 1 abortion 4 or 1 his 2 political 1 ambitions 4 ; 6 the 1 Massachusetts 3 Supreme 1 court 4 has 1 ruled 4- that 2 If 1- a 1 fetus 2 is 1 viable 4 ; 5 it's 2 considered 1- a 1 person 4 under 1 the 1 motor 1- vehicle 1 homicide 1 law 4 ; 6 and 1 Harshbarger 3- says 4 anyone 4 who 1 kills 2 a 1 viable 1 fetus 4 by 2 driving 1 drunk 4 ; 5 ought 2 to 1 be 1 prosecuted 4 ; 6 Levy 3 is 1 not 1 commenting 4 but 1 her 1 attorney 4- Nancy 1 Gertner 4 ; 5 says 3 the 1- homicide 3- charge 4 is 1 absurd 4 ; 6 Gertner 1 says 4 the 1 state 2 Supreme 1 Court's 1 ruling 4 ; 5 involved 2 the 1- case 3 of 1- a 2 third 2 party 4 whose 1 car 3 hit 1- a 1 pregnant 1 pedestrian 4 killing 1 her 1 fetus 4 ; 6 but 1- according 2 to 1- Gertner 4 the 1 high 1 court 2 said 2 nothing 4- about 2 prosecuting 1 a 1 woman 4 for 1 harming 2 her 1 own 2 fetus 4 ; 6 It 1 may 1- be 2 appropriate 4 to 1- charge 1 Levy 3 with 2 drunk 1 driving 4 says 1 Gertner 4 ; 5 but 0 it's 2 unconstitutional 4 to 1 also 1 charge 1 her 4 ; 5 with 1 killing 4- part 1- of 1 her 1 own 2 body 4 ; 6 Director 3- of 1- the 1 Massachusetts 2 Civil 1 Liberties 1 Union 4 John 1 Roberts 4 ; 5 wants 1 Harshbarger 3 to 1 drop 2 the 1 homicide 1 charge 3 Against 1 Levy 4 ; 6 the 1- case 3 is 1 part 2 of 1- an 1- alarming 1 trend 4 he 1 says 4 ; 5 to 1- pit 2 so-called 2 fetal 1 rights 4 Against 2 women's 1 rights 4 to 1- privacy 3 and 1 autonomy 4 ; 6 but 1 Harshbarger 3 says 4 ; 5 there's 1 no 1 threat 4 that 2 pregnant 1 women 3 will 1 be 1 prosecuted 3- for 2 legal 1 conduct 4 ; 5 even 1- If 4 harm 2 to 0 a 1 fetus 3 results 4 ; 6 Marian 2 Stokes 1 agrees 4 saying 3 that 2- after 1 all 4 ; 5 drunk 2 driving 4 is 1- Against 2 the 1 law 4 ; 6 the 1 co-founder 3- of 0 the 1 greater 1 Boston 1 chapter 3 of 2 Mothers 1 Against 1 drunk 1 driving 4 ; 5 says 3 she 1 empathizes 4 with 1- the 1- pain 3 Elizabeth 1 Levy 3- must 1- be 1 feeling 4- over 1 her 1 loss 4 ; 5 but 0 she 1 says 4 that's 2 no 1 reason 3 to 0 treat 2 Levy 4- differently 4 ; 5 than 1 the 1 drunk 1 driver 3 who 1 killed 3 Stokes' 1 son 4 ; 6 If 2 the 1 motor 1- vehicle 1 homicide 1 charge 2 sticks 4 and 1-

Levy 1 is 1 convicted 4 ; 5 she 1 faces 3- one 3 to 1- fifteen 1 years 4 in 1 prison 4 ; 6 for 2- WBUR 4 I'm 1- Margo 1 Melnicove 4 ; 6

STORY 33

Like 2 many 1 Rockport 1 residents 4 ; 5 Martha 1 Woodworth 4 is 2 furious 3 with 2- the 1 Smithsonian 4 ; 6 Woodworth 4 an 1 astrologer 2- and 1- psychic 4 ; 5 says 3 she 1 hasn't 1 been 1- in 1- touch 2 with 2 Frans 1 Denghausen 4 ; 5 a 1 longtime 2 Rockport 1 artist 4 who 1 died 3 three 1 years 1 ago 4 ; 6 but 1 she's 1 certain 3 Denghausen 4- would 1 be 1 furious 2 too 4 ; 6 At 0 issue 4- is 1 a 1 codicil 2 to 1- his 1 will 4 ; 5 bequeathing 3 one 1 million 1 dollars 2 to 1 Rockport 4 to 1 replace 3 it's 2 obsolete 2 public 1 library 4 ; 5 provided 2 the 1 town 3 contributes 1 land 4 for 1 a 1 new 1 building 4 ; 5 or 1 funds 3 to 1 renovate 4 an 1 existing 1 structure 4 ; 6 the 1 Smithsonian 4 the 1 beneficiary 3 of 2 the 1 bulk 2 of 1 Denghausen's 1 estate 4 ; 5 claims 3 the 1- bequest 3 can 1 be 1 used 4 only 2 for 1 library 2 operating 1 costs 4 ; 5 once 1 Rockport 2 uses 1 it's 1 own 1 money 4 for 1 capital 1 improvements 4 ; 6 If 1 Rockport 3 does 2 not 1 inherit 3 the 1 million 1 dollars 4 ; 5 the 1 Smithsonian 4 as 2 residual 1 beneficiary 4 will 4 ; 6 Admittedly 4- the 1 codicil 4- is 2 confusing 4 ; 5 but 1 Denghausen's 2 friend 4 and 1 former 1 library 1 trustee 4 Anne 1 Fisk 4 ; 5 says 4 in 1 several 1 conversations 4 before 1 his 2 untimely 1 death 4 ; 5 Denghausen 2 indicated 4- he 1 wanted 1 the 1 money 4- to 1 be 1 spent 4 on 1 converting 4- a 1 Now 1 closed 2 school 4 into 1 the 1 new 1 library 4 ; 6 town 1 officials 1 say 4 given 2 that 1 Rockport's 2 already 3 heavily 1 in 1 debt 4 to 1- pay 3 for 1- a 1 new 2 school 1 complex 4 ; 5 there's 1 no 2 new 1 library 3 in 1 sight 4 unless 2 Denghausen's 1 money 4 can 1- be 1 used 4- for 1 construction 4 ; 6 the 1 present 1 library 4- opened 4- in 1 nineteen 1- oh 1 six 4 ; 5 director 3 Steven 1 Rask 4 apologizes 4 for 1 piles 2 of 1 books 2 on 1 the 1 floor 4 ; 5 saying 3 there's 2 no 1 more 1 room 2 on 1 the 1 shelves 4 ; 6 If 1- allowed 2 to 1 move 3 to 1 the 1 former 2 school 1 house 4 Rask 2 says 3 the 1 library 4- could 1 accommodate 4 nearly 2 three 1 times 1 as 1 many 1 volumes 4 ; 5 an 1 expanded 2 children's 1 room 4 a 1 community 1 meeting 1 hall 4 and 0 a 1 reading 1 room 4- named 1 after 2 its 1 benefactor 4 ; 6 Now 4- there's 2 no 1 place 1 to 1- read 4 says 1 Rask 4 ; 5 except 2 for 1- a 1 few 1 chairs 4 squeezed 2 into 1 corners 4 ; 6 Smithsonian 1 spokesperson 4 Madeleine 1 Jacobs 4 ; 5 insists 3 the 1 institution's 4 not 2 Being 1 greedy 4 ; 6 the 1 Smithsonian 3 already 1 inherited 4 nearly 2 four 1 million 1 dollars 4 from 1 Mrs 1 Denghausen 4 who 1 died 2 a 1 year 2 before 1- her 1 husband 4 ; 6 it's 1 due 3 four 1- to 1- five 1 million 1 more 4 once 2 Mister 1 Denghausen's 1 estate 4 is 1 settled 4 ; 6 Being 1 entrusted 3 with 2 most 1 of 1 the

1 couple's 1 fortune 4 says 1 Jacobs 4 ; 5 gives 3 the 1 Smithsonian 4 a 1 responsibility 4 to 1 ensure 4 that 2 none 1- of 1 it's 4 misspent 4 ; 6 Rockport 1 residents 4 find 2 that 1 argument 4 especially 1 galling 4 ; 5 saying 3 to 2 the 1 Smithsonian 4 Denghausen's 2 nothing 1 more 4- than 1- a 1 generous 1 stranger 4 ; 6 the 1 case 2 will 1 be 1 heard 3 in 1 Essex's 2 probate 1 court 2 August 1 fifteenth 4 ; 5 meanwhile 4 several 2 of 1 Rockport's 3 well 1 connected 1 citizens 4 ; 5 are 1 cancelling 4- their 2 Smithsonian 1 memberships 4 ; 5 and 1 lobbying 1 politicians 4 to 1- throw 1 their 1 weight 4 behind 1 the 1 town 4 ; 6 If 1 Rockport 1 prevails 4 ; 5 the 1 Smithsonian's 2 Jacobs 1 says 4 the 1 institution's 3 likely 2 to 1- let 2 the 1 matter 4 rest 1 in 1 peace 4 ; 6 for 1 WBUR 4 I'm 1- Margo 1 Melnicove 4 ; 6

STORY 34

Wakefield 1 businessman 4- Raymond 1 Marshall 4 ; 5 says 2 his 1 property 3 on 1 the 1 edge 2- of 1 Grafton 4 is 1 an 1 ideal 1 location 4 for 1 the 1 asbestos 2 Conversion 1 plant 4 ; 6 Marshall 1 says 4 it 1- would 1 take 2 about 2 ten 1 million 1 dollars 4 to 1 equip 3 the 1 building 4 for 2 glass 1 making 4 ; 5 with 2 six 1 furnaces 3 operating 2 round 1 the 1 clock 4 ; 5 to 1 melt 1 down 3 truckloads 2 of 1 asbestos 1 waste 4 each 1 day 4 ; 6 Marshall 1 says 2 the 1 venture 3 would 1 be 1 good 2 for 1 Grafton 4 ; 5 asbestos 1 haulers 4 would 1 pay 1 a 1 fee 2 for 1 disposal 4 ; 5 and 2 Marshall's 1 agreed 3 to 0 give 2 ten 1 percent 3 to 1 the 1 Town 4 ; 5 which 1 he 1 says 2 could 1- amount 2 to 1 more 1 than 3 three 1 million 1 dollars 3 annually 4 ; 5 nearly 2 one 1 third 3 of 1 Grafton's 2 total 1 budget 4 ; 6 the 1 operation 3 would 1 also 1 be 1 good 4 for 2 Marshall's 1 profit 1 margin 4 ; 5 although 3 he 1 refuses 1 to 1- be 2 specific 4 ; 6 and 1- Marshall 1 says 3 it 1 would 1 be 1 good 2 for 1 the 1 environment 4 to 1- have 2 cancer 1 causing 3 asbestos 1 fibers 4 ; 5 removed 2 from 1- it 4 forever 4 ; 6 once 3 converted 2 from 1 asbestos 4- into 1 a 1 low 1 quality 3 dark 1 colored 1 glass 4 ; 5 Marshall 1 says 2 the 1 material 4 could 1- be 1 used 4 as 1- a 2 gravel 1 replacement 4 which 4 when 2 mixed 2 with 1 asphalt 4 ; 5 provides 3 a 1 road 1 surface 3 with 2 superior 1 grip 4 ; 6 if 1 there 2- are 1 no 1 takers 4 ; 5 Marshall 1 says 3 the 1 glass 1 pellets 4 would 1 merely 1 become 2 non-toxic 1 landfill 4 ; 6 But 2 University 1 of 1 Lowell 3 Occupational 1 Health 1 Professor 4 Raphael 1 Maurri 4 ; 5 says 3 once 3 asbestos 2 is 1 removed 3 from 1 a 1 building 4 then 2 wetted 3 and 1 bagged 4 according 2 to 1 government 1 regulations 4 ; 5 it's 2 no 1 longer 3 dangerous 4 ; 6 Maurri 1 says 3 it 1 makes 2 no 1 sense 4 to 1- disturb 3 the 1 material 3- again 4 ; 5 for 1 the 1 sake 3 of 1 recycling 4 ; 6 Grafton 1 residents 2 say 4 they're 1 worried 3 about 2 what 1- could 1- happen 4 even 2 before 2 the 1 material 4- gets 2 to 1- the 1 plant 4 ; 6 En 1

route 4 asbestos 1 filled 1 trucks 4 would 1 have 1- to 1- drive 3 through 1 residential 2 areas 4 ; 6 Only 2 one 1 accident 4 says 2 the 1 town's 3 Public 1 Health 1 nurse 4 ; 5 could 1 subject 2 countless 1 homes 4 to 1- an 1 asbestos 1 shower 4 ; 6 Town 1 planner 4 Peter 1 Lowett 2 says 4 more 1 research 2 is 1 needed 4 to 1 determine 3- if 2 that 1 fear 3- is 1 well-founded 4 ; 6 Meanwhile 4 he 1 says 3 Town 1 officials 2 will 1 not 2 rubber-stamp 1 the 1 project 4 ; 5 even 1 though 2 Grafton 4 could 2 certainly 1 use 4 the 1 millions 1 of 1 dollars 4 it's 1 been 1 promised 4 ; 6 Beyond 2 Town 1 zoning 3 and 2 Public 1 Health 1 Board's 1 review 4 ; 5 a 1 state 1 environmental 1 impact 1 report 4- is 1 likely 2 to 1 be 1 required 4 ; 6 Developer 2 Raymond 1 Marshall 1 says 4 if 2 the 1 permitting 1 process 3 drags 1 on 4 he'll 1 take 1- his 1 business 4 out 1 of 1 state 4 for 1 WBUR 4 I'm 1- Margo 1 Melnicove 4 ; 6

Bibliography

- Chen, M. (1985), *The syntax of Phonology: Xiamen tone sandhi*, University of California, San Diego.
- Chomsky, N. (1995), Bare phrase structure, in G. Webelbuth, ed., ‘Government and binding theory and the minimalist program’, Blackwell publishers, pp. 383–439.
- Collier, R., de Pijper, J. & Sanderman, A. (1993), Perceived prosodic boundaries, in L. Bates, ed., ‘Proceedings of the ARPA Human Language Technology Workshop’.
- Firbas, J. (1964), ‘On defining the theme in functional sentence analysis’, *Travaux Linguistiques de Prague* **1**, 267–280.
- Fujisaki, H., Hirose, K., Takahachi, N. & Morikawa, H. (1986), Acoustic characteristics and underlying rules of intonation of the common Japanese used by radio and television announcers, in ‘Proceedings of the International Conference of Acoustics, Speech and Signal Processing’, pp. 2039–2042.
- Gee, J. & Grosjean, F. (1983), ‘Performance structures: A psycholinguistic and linguistic appraisal’, *Cognitive Psychology* **15**, 411–458.
- Grimshaw, J. (1990), *Argument Structure*, Cambridge, Mass; London: MIT Press.
- Gussenhoven, C. (1983), *On the Grammar and Semantics of Sentence Accent*, Dordrecht, Foris.

- Hahn, U. (1995), Distributed text structure parsing-computing thematic progression in expository texts, in G. Rickheit & C. Habel, eds, 'Focus and Coherence in Discourse Processing', Berlin:Walter de Gruyter, pp. 214–250.
- Hajicova, E., Skoumalova, A. & Sgall, P. (1995), 'An automatic procedure for topic-focus identification', *Computational Linguistics* .
- Hale, M. & Selkirk, E. (1987), 'Government and tonal phrasing in papago', *Phonology Yearbook* **4**, 151–183.
- Halliday, M. (1967), *Intonation and Grammar in British English*, Mouton, The Hague.
- Hayes, B. (1995), *Metricall stress theory:Principles and case studies*, Chicago:University of Chicago Press.
- Hoffman, B. (1996), Translating into free word order languages, in 'Proceedings of the International Conference onComputational Linguistics', pp. 556–561.
- Inkelas, S. (1989), Prosodic constituency in the lexicon, PhD thesis, Stanford University, Palo Alto, CA.
- Jackendoff, R. (1972), *Semantic interpretation in Generative Grammar*, MA:MIT Press, Cambridge.
- Komagata, N. (1998), Information Structure in Texts: A Computational Analysis of Contextual Appropriatness in English and Japanese, PhD thesis, University of Pennsylvania.
- Kruijff-Korbayova, I. & Webber, B. (2001), Concession, implicature and alternative sets, in 'Proceedings of the International Workshop on Computational Semantics', pp. 227–248.
- Ladd, D. (1986), 'Intonational phrasing: The case of recursive prosodic structure', *Phonology* **3**, 311–340.
- Ladd, D. (1996), *Intonational Phonology*, Cambridge:Cambridge University Press.

- Marantz, A. (1995), The minimalist program, in G. Webelbuth, ed., 'Government and binding theory and the minimalist program', Blackwell publishers, pp. 349–382.
- McCarthy, J. & Prince, A. (1993), 'Generalized alignment', *Yearbook of morphology* pp. 79–153.
- Nespor, M. & Vogel, I. (1986), *Prosodic Phonology*, Dordrecht:Kluwert.
- Ostendorf, M., Price, P. & Shattuck-Hufnagel, S. (1995), 'The boston university radio news corpus'.
- Partee, B. (1996), Allegation and local accomodation, in B. Partee & P. Sgall, eds, 'Discourse and Meaning', Amsterdam:John Benjamins, pp. 65–86.
- Peperkamp, S. (1997), Prosodic words, PhD thesis, Universiteit van Amsterdam.
- Pierrehumbert, J. (1980), The phonology and phonetics of English intonation, PhD thesis, MIT.
- Pierrehumbert, J. & Beckman, M. B. (1988), *Japanese Tone Structure*, Cambridge, Mass: MIT Press.
- Pierrehumbert, J. & Hirschberg, J. (1990), The meaning of intonational contours in the interpretation of discourse, in P. Cohen, J. Morgan & M. Pollack, eds, 'Intentions in communication', MA:MIT Press.
- Prevost, S. & Steedman, M. (1996), Generating contextually appropriate information, in 'Proceedings of the Conference of European Chapter of the Association for Computational Linguistics', Vol. 34, pp. 294–301.
- Price, P., Ostendorf, M., Shattuck-Hufnagel, S. & Veilleux, N. (1988), 'A methodology for analyzing prosody', *Journal of the Acoustical Society of America* **84**.
- Prince, E. (1981), Toward a taxonomy of given-new information, in 'Radical Pragmatics', New York:Academic Press, pp. 223–256.
- Quirk, R., Greenbaum, S., Leech, G. & Svartnik, J. (1985), *A Comprehensive Grammar of the English Language*, Longman Inc., New York.

- Selkirk, E. (1978), On prosodic structure and its relation to syntactic structure, in T. Fretheim, ed., 'Nordic Prosody II', Trondheim:TAPIR.
- Selkirk, E. (1984), *Phonology and Syntax. The Relation between Sound and Structure*, Cambridge, Mass: MIT Press.
- Selkirk, E. (1986), 'On derived domains in sentence phonology', *Phonology Yearbook* **3**, 371–405.
- Selkirk, E. (1993), Constraints on prosodic structure-a modular approach, in 'ESCA Workshop on Prosody'.
- Selkirk, E. (1995), 'The prosodic structure of function words', *University of Massachusetts occasional papers 18: Papers in Optimality Theory* pp. 439–469.
- Shattuck-Hufnagel, S. & Turk, A. (1996), 'A prosody tutorial for investigators of auditory sentence processing', *Journal of Psycholinguistic Research* **25**(2), 193–246.
- Silverman, K., Beckman, M., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., Pierrehumbert, J. & Hirschberg, J. (1992), Tobi: a standard for labeling english prosody, in 'Proceedings, Second International Conference on Spoken Language Processing', Vol. 2, Banff, Canada.
- Steedman, M. (2000a), 'Information structure and the syntax-phonology interface', *Linguistic Inquiry* **31**, 649–689.
- Steedman, M. (2000b), *The syntactic Process*, MIT Press, Cambridge MA.
- Taylor, P. (2000), 'Analysis and synthesis of intonation using the tilt model', *Journal of the Acoustical Society of America* **107**, 1697–1714.
- Truckenbrodt, H. (1999), 'On the relation between syntactic phrases and phonological phrases', *Linguistic Inquiry* **30**(2), 219–255.
- Wightman, C., 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**(3), 1707–1717.