PROBLEMS OF INTONATION

Mary E. Beckman
Department of Linguistics, Ohio State University, Columbus, OH, USA
and ATR Interpreting Telecommunications Research Laboratories, Kyoto, Japan

DELIMITING THE PROBLEM

My advertised role in this symposium is to talk about "problems of intonation" in teaching phonetics. The first step is to delimit what we mean by intonation, and to understand why there are problems. So, to repeat my earlier definition, let me delimit the object of study as: "**all** aspects of the perceived pitch pattern that the speaker intends for the hearer to use in understanding the utterance, or that the hearer does use whether intentionally controlled by the speaker or not" [1]. Intonation is one of the most difficult aspects of speech to teach about, for the same reasons that it is one of the most difficult aspects of language sound structure to model.

The problem of meaning

One difficulty stems from the kinds of meaning that many intonational categories have, which are like the meanings of such words as *the* versus *a* in English, or the meaning of choosing the indicative versus subjunctive form in Portugese. They can signal often very subtle facts about the relationship between an utterance and its larger discourse context.

For example, in the American English intonation system, there is a categorical contrast between high tone (H*) and low tone (L*) pitch accents which in different discourse contexts can be interpreted as the difference between a statement and a yes-no question, between an imparting of new information and a gentle reminder of old information, between an affirmative repeating and an incredulous echoing of what the other speaker has just said, or between the literal use of a word such as *now* and its use as a discourse marker for a shift in discourse topic. What do all these usages have in common? Pierrehumbert and Hirschberg [2] have suggested that H* means commitment — the speaker intends the hearer to add a particular proposition or entity to the background of mutually believed information at that point in the discourse — whereas L* means lack of commitment, either because the speaker intends the hearer to provide the correct information (yes-no question), or because the speaker knows that the hearer already has the information (gentle reminder), or because the speaker questions the other speaker's intent to add the information (incredulous echo), or because the speaker intends that the word should not be taken as part of the substantive information content of the narrative (discourse marker usage).

This is a meaning difference that is inherently more difficult to characterize than many of the lexical differences that we use to support our analyses of the consonants and vowels of a language. A speaker of English need only point to exemplar objects to characterize the meanings of *mat* and *bat*, and thus get across that these are different morphemes. And it is not difficult to find a slew of other morphemes such as *mitt* versus *bit* or *moat* versus *boat* to support an analysis that localizes the difference to the beginnings of the morphemes in order to claim that [m] and [b] are discretely different consonant sounds that are differentiated in onset position in English.

By contrast, there is fervent argument among experts concerning the analogous questions in the analysis of the intonational morphology of English, and even those ultimate experts — children acquiring English as their native language — apparently find these kinds of meaning differences more difficult to learn to control (see [3], [4]). It is hard to get the phonetic skills relevant to studying intonation without also giving oneself at least a nodding acquaintance with the related literature in semantics, pragmatics, and discourse structure.

The alphabetic (dis)advantage

Intonation is difficult to teach also because our modeling of it begins from scratch. We do not start with an already highly practiced phonological theory, as we do when teaching our students how to analyze consonant and vowel contrasts.

Perhaps because the orthographies of English, French, and German are conservative, we tend to concentrate on the difficulties that our students' literacy imposes, to bemoan how difficult it is to get students to transcribe a form how they hear it and not how they know it is spelled. However, research on speakers who have learned only a logographic writing system [5] (or who find it difficult to learn to read an alphabetic one [6]) makes it clear that learning to read in an alphabetic writing system imparts (or requires) a very sophisticated meta-linguistic phonological awareness, an awareness that we take full advantage of in our standard method for teaching about vowel and consonant contrasts. We simply assume the linear segmental phonological analysis of the International Phonetic Alphabet, and concentrate on teaching other things — on developing the ability to hear and mimic consonant and vowel contrasts not in our students' native repertoires, and on imparting basic facts and physical principles that our students need to know in order to understand how the contrasts are produced and perceived.

However, this method is not without cost. Our blithe assumption of the alphabetic model creates the Platonist illusion that consonants and vowels exist in nature already segmented that way, that there are countable "sounds" out there for the phonetician to study independent of the phonological analysis. We imagine that we can get all the skills that we need in order to know about how structure is transmitted in the speech stream without developing our ability to imagine and evaluate alternative models of the structure transmitted. This self-delusion then handicaps us when we try to teach about "suprasegmentals" — about any feature of language sound structure that defies this particular phonological model, as intonation so clearly does. So we teach our students about the physiology of the larynx and the physical relationship between vocal fold tension and fundamental frequency, and we are baffled that it is so difficult to make them understand the dimensions of intonational variation available to them in their native languages, or to seewhy ostensibly the same pitch contour can function so differently in different parts of an utterance or in utterances from another language.

The difficulty of prosody

This difficulty is compounded by the fact that so many aspects of intonation are crucially linked to prosodic constructs such as the syllable, stress, and phrasing. For example, we cannot explore such intonational constructs as pitch accents and boundary tones without recourse to prosodic units such as the syllable and the intonational phrase, and we cannot explore the different functions of pitch accents in English versus Japanese without recourse to the prosodic construct of stress. These prosodic constructs are themselves difficult to model for our students, again because they do not fit the alphabetic model, but also for a more fundamental reason. I think prosody is difficult by its very nature, because prosody is all about the organizational structure of speech (as opposed to its contrastive structure), and I think organizational features are harder for the human mind to grasp.

Let me try to explain this idea with a metaphor. Think of a jacquard weave in which the design is made not just by varying whether the warp or the woof is the surface thread, but also by varying the thickness along the length of each thread and the saturation and color of the dye. These are all features that can be modeled to some extent without specifying how any particular variation functions at any place in the weave. To predict the effect of some change in the dye, for example, the expert weavers need to know about the physical characteristics of the dye and how color is perceived and so on. But now consider the shapes that the weavers are trying to build in the surface of the cloth. To model these, the weavers will need to know about such things as edge detection and so on, but also they need to specify what aspects of the thread they can vary at any place in the weave to define the desired edge.

Features of vowel timbre, consonant place and manner, phonation type, pitch level, and so on, are easier because they are all thread features. We can talk about them to a first approximation without thinking about how they function in any particular part of the organizational structure of a given language. For

example, we can talk about segmental effects on the pitch contour and the psychoacoustics of pitch perception and even the categorical intonational features H ("high tone") versus L ("low tone") without specifying whether these features function to differentiate monosyllabic verb morphemes (as in Yoruba) or differentiate single-tone pitch accents (as in English).

Prosodic constructs such as the syllable, stress, and phrasing, on the other hand, are like the shapes in the cloth that the weavers make by alternating woof and warp, thick and thin, or dark and light colored thread. They are features of the organization itself which depend crucially on what thread features can alternate where. They are about that difficult question precluded by the Platonist illusion: What are the possible units of counting for a particular language? What are the segmentations that the speaker can intend for the listener to parse in the signal?

OVERCOMING THE PROBLEM

Of course many other aspects of language sound structure are also very difficult. The standard acoustic model of vowel contrasts is extremely difficult for the unfortunately large number of our students who lack the background in physics and the necessary mathematical skills to grasp the basic principles of resonance. But (as Vassiere points out in her paper) there are dramatic examples — such as the film clip of the Tacoma Narrows Bridge bouncing rhythmically in the breeze as it gets ready to break apart — to give an initial intuitive feeling for the phenomenon, which then can be used to motivate a more precise handwaving.

An important element to success in teaching the skills relevant to intonation, then, is to exercise just this kind of showmanship. In the rest of this paper, then, I will try to give a few of the examples from English and other languages that have worked for me in developing strategies for overcoming each of the three main sources of difficulty for intonation.

Start with some easier meanings

One common strategy for overcoming the problem of meaning is to first teach about a language in which different tone patterns differentiate ordinary lexemes, before teaching about the difficult-to-characterize pragmatic morphemes that tone patterns differentiate in English, German, or Italian. If the language chosen for this purpose has an intonational system such as that of Swedish [7] or Osaka Japanese [8], a system which allows a relatively easy progression from hearing the differences in short citation forms to decomposing longer forms into aspects that are part of the lexical specification of the words and aspects marking more elusive discourse functions, this method can provide a more gentle introduction into the terrain. It allows the students to concentrate first on some of the other skills that they will need in order to overcome the alphabetic disadvantage, before venturing into the crags and crevices of pragmatics and discourse structure. The method only works, however, if the teacher is a fluent speaker of the language, or if the teacher can bring into the classroom native speaker consultants who have strong intuitions about pragmatic felicity and about any relevant prosodic constructs.

The alternative strategy is to choose to teach first the intonational system of a language which all of the students must know very well to be in the class — the language of instruction — and to choose the initial examples to be as vivid and salient as possible. In some places, such as the linguistics department of the Ohio State University, where I teach, many of the students will even be native speakers of the language. The first examples then can be situations in which failing to differentiate two intonation patterns can have an embarrasingly funny effect, as in the two utterances:

(1) Mary does intonation.
L+H* L- L%

(2) Mary does intonation.
L*+H L- H%

which I have used to illustrate the difference between the L+H* and L*+H accents in English. I offer the two as alternative responses to the claim that "Only crazy people do intonation." The students in the class who are native speakers of English laugh at the second response (which implies that the teacher is crazy), and those who are not become very motivated to learn to control the contrast between the two tunes.

Sometimes the most salient examples involve a difference in interpretation that might be signaled also by a syntactic difference. The L* H- H% pattern in American English, for example, is often called the "yes-no question contour" in differentiation to the "declarative contour" H* L- L%. It is convenient to use a yes-no question and a declarative sentence with these tunes, in order to build two parallel series of utterances showing how the H- versus L- phrase accents extend over longer and longer regions as the nuclear pitch accent is moved earlier and earlier. However, the teacher must be careful also to provide examples of utterances where H* L- L% occurs on a yes-no question and where L* H- H% is used as a statement (e.g. the statement of incredulous disbelief), lest some students get the mistaken impression that some tunes are "syntactic" and therefore more worthy of their attention (or less worthy, depending on their feelings about areas of grammar other than phonology and phonetics). The example I like to use is the joke about the man forced to read a public confession, who renders the confession with great feeling, dividing he sentences into many intonational phrases:

(3) I || was wrong, || and Stalin || was right. || I || should be vilified. || ...

and putting a L* H- H% pattern on each. It will be obvious from the context of the story that the man cannot be asking his audience a series of yes-no questions.

Because jokes and funny stories in general have this advantage of building an explicit fixed context into the performance of the example, they can also be used to help get across an important lesson. In order to investigate intonational form, one must carefully observe (or even actively manipulate) the discourse context for the speaker. As part of inculcating good observational habits, students will need to be debunked of the notion that asking the speaker to recite a particular sentence or phrase for the field linguist somehow provides a "neutral" context, and that the prosodic organizations and associated intonation contours produced in this context have special status as "neutral" or "default" patterns. We need to make it clear that all this method accomplishes is to make it impossible to observe the context which the speaker then implicitly imagines in order to be able to recite the form. If this makes the students feel a bit daunted at the prospect of having to learn how to judge discourse contexts as well as to hear tunes and stresses and phrasings, better to have them feel daunted than to send them into the field thinking that they will be able to model intonation in a new language without acquiring all of the requisite observational skills.

Use the F0 contour as the narrow phonetic representation

When teaching an introductory general course in phonetics, one may be tempted to try to shoehorn intonation into the alphabetic model by teaching the IPA symbols for "high tone", "falling word accent", "global rise", "downstep", and so on. I think this is a mistake.

Of course, students specializing in a particular language or language area may want to learn some of these symbols in order to have access to the literature and fieldwork notes of other linguists. For example, students specializing in Chinese or Tibeto-Burman linguistics may want to learn the nicely iconic Chao tone letters [9], which are the basis for the current IPA symbols for "tones & word accents". On the other hand, they will also need to learn the less iconic superscript numbers transliterating the tone letters, since this is the transcription that they will encounter in reading this literature, and students specializing in Bantu languages will need to learn the even less iconic non-IPA diacritics that many Africanists use [10]. But such specialist reading knowledge is quite different from the phonetic skills that one needs to teach to enable the students to observe intonation patterns and analyze intonational systems *in vivo*. Since the segmentations imposed by the alphabetic analysis are so wrong for intonation, I think it is better to avoid any symbolic transcription until one is confident of the phonological categories. I think it is far better to sidestep the IPA entirely and expend the same effort on teaching students to look at raw F0 contours while making stylized drawings of the pitch contours that they hear. In other words, we should train the students to use the F0 contour in lieu of an initial "narrow transcription" in doing fieldwork.

Since this may be construed as a radically anti-IPA stance, let me explain it further by considering where the narrow transcription fits in a standard strategy for teaching observational skills relevant to analyzing consonant and vowel systems, the strategy that structures the sequence of chapters and topics in Peter Ladefoged's textbook [11], which we use in our introductory course at Ohio State University. The first step in the sequence is analogous to the second strategy I suggested above for overcoming the problem of intonational meaning. We begin by teaching the students to transcribe the consonant and vowel phonemes of the language of instruction, which for many students will be the native language. Then we make them produce narrow transcriptions of that language, before we go on to teach them about other languages. We use that intervening step of narrow transcription to reduce the phonological interference from the first language that will otherwise hinder their observations when they go to do fieldwork. We do this because it is far less threatening than simply putting them into the field, where they would have to worry about the meanings of forms and a host of other things (such as the cultural norms about eye contact and so on), at the same time as they are trying to learn to hear contrasts not found in the native language. In other words, by treating allophonic variation in consonant and vowel quality in the native language as if it were alphabetic contrast, we give the students a Zen exercise to help them turn off the attentional skills built up over a lifetime of learning to perceive the phonemes of the native language. This works because we assume that for any language they will encounter in the field, it would not be disastrously wrong to posit as a first working hypothesis, a linear segmental analysis of the language's system of consonant and vowel contrasts.

Applying the same strategy to the analysis of intonation is a recipe for making blindfolds. It may be safe to assume that an alphabetic tone-sequence analysis will work for any new intonation system encountered in the field, but *a priori* we cannot assume any of the other facts we need to have in order to uncover the tonal analysis, answers to questions such as: Do the different tones contrast paradigmatically in composing pragmatic morphemes (as in English) or lexical specifications (as in Cantonese), or is their primary role the syntagmatic marking of prosodic constructs such as stress (as in Danish) and accentual phrase edges (as in most dialects of Korean)? What are the relevant prosodic constructs for anchoring the tones to the consonant and vowel features in an utterance, and are there contrasts in temporal alignment between the tones and the anchor site? How do these facts about the relationship between tone contrasts and meanings and about the anchoring of tones to other features of contrastive structure translate into density of tonal specification? Are some tone sequences or some prosodic constructs associated with a systematic manipulation of the more global pitch pattern, such as downstep or pitch range reset, which can obscure the local tonal values? Intonation systems can differ so markedly along all of these dimensions that prematurely adopting any symbolic transcription obstructs the observation of variation in pitch necessary to getting the phonological analysis. To give just one example, if Janet Pierrehumbert and I had not decided to use a nonsymbolic phonetic representation, the F0 contour, we would not have been able to observe the relationship between F0 slope and phrase length, and so on, observations that challenged the traditional narrow phonetic transcription with its mora-by-mora specification of high versus low tone and suggested the analysis of the Japanese intonation system proposed in [12].

This is not to say that I do not advocate teaching symbolic labels for intonational categories. Far from it. I use the ToBI labels H*, L*, L+H*, and so on, every time I teach about English intonation, because the literature from the past half century of work on the intonational categories and meanings of American and Southern British English makes me confident that this analysis will work for the dialects spoken by the majority of my students. My confidence is bolstered by the high degree of intertranscriber agreement documented for this transcription system [13], and by the fact that the system has been taught successfully to non-native speakers

employed to label spontaneous dialogue [14] and has even been used by a non-native speaker to teach about English in a graduate phonetics seminar on intonation [15]. A phonetician teaching intonation to a class of Swedish speakers or Japanese speakers or Spanish speakers or Mandarin Chinese speakers similarly would be justified in teaching a symbolic transcription of the intonation patterns, because we know what the categories are. But when there is no established body of research on intonational form and meaning for a language, we would do better not to use that language as the obvious first extended example even if all the sudents in the class are native speakers. In such a case, we might teach the students first about a language that does have an established phonological analysis. We might give them the necessary fieldwork skills to control the pragmatic context in eliciting forms and to interpret the F0 track, and encourage them to embark on the research toward such an analysis for their own language. But we should not fudge by offering them a symbolic "narrow phonetic transcription" of intonation to stand in until the real work is done.

What skills do our students need in order to use the F0 track in lieu of a narrow transcription? The most important thing at first is that they will need to learn about microprosody. One might be tempted to "doctor" F0 contours for the students at early stages by digitally smoothing and manually whiting out any apparent "perturbation" remaining after the smoothing, but I think this is a mistake. No non-intelligent smoothing procedure can do what the human mind does, and the students need to train their eyes to do what their ears' minds do: to parse over the effects on intonation of consonant and vowel features when they are "listening" for tonal contrasts. We do better to start this training immediately, by giving them examples of the same tone pattern anchored to different consonants and vowels. When they have mastered the art of visually "listening" through the microprosody, then we can tell them about the places where F0 is not a good measure of pitch. They need to know to look for the manifestation of vocal fry and other non-modal phonation types. At this point it is good to set them to designing little experiments. We can ask them how they might go about comparing slopes of two different kinds of F0 rise in the language being studied, or how they might test competing hyptheses about the timing of a tone relative to some prosodic landmark.

Grapple with prosody

Which brings us to the most difficult area of all. There is no way to tackle intonational analysis without grappling with the problem of prosody. We must teach our students about the nature of prosody and give them the skills to uncover the prosodic constructs that native speakers of a language use to structure the intonational categories and other contrastive features.

Again, the best strategy here follows the same lines as the best strategy for teaching about tone features. We can start with a language that all of the students know fluently, where we can tap their intuitions about relevant prosodic constructs such as how many syllables are in a word or how many intonational phrases are in a longer utterance. For each prosodic construct, it is good to start with clear cases, where native speakers (and linguists) tend to agree on the number or location of relevant prosodic landmarks, before proceeding to cases where the analysis is more unsure or simply unknowable. For example, in teaching about stress in English, it is best not to start with words such as *thirteen* and *Malay*, where different students may produce different nuclear accent placements in the citation form, and which would lead us directly into the nettles patch of stress shift. (Note that these are the words that some pronunciation dictionaries transcribe as having two primary stresses.) On the other hand, we also should not gloss over the difficulty of prosody, but attempt from the beginning to get across the idea that it prosody about organizational structure. It is a mistake to think we can make it easier for them by first offering misleading definitions in terms of contrastive structure. For example, we are unwise to say: "Just as tone can be defined in terms of relative pitch, stress is the relative loudness of a syllable." A stressed syllable may indeed sound louder than an unstressed syllable at any

level of the stress hierarchy, but this phonetic equation will lead the students straight down the dead-end track of looking for stress in the RMS amplitude contour.

I find that it is easiest to grapple with prosody if I first cover the phenomena of "coarticulation". That is, I place prosody in the pedagogical sequence where I can more easily show how prosodic structure is realized substantively by often very subtle differences in coordination among different contrastive features. I then can relate questions about prosodic constructs to questions that have already come up in the discussion of segmental features and coarticulation, such as: "What is the difference between an affricate and a cluster of stop plus fricative, anyway, and why isn't the [ts] at the end of *cats* an affricate?" In answering these questions I will already have stressed a point that is crucial for understanding higher levels of structural organization — namely, that what we transcribe as ostensibly the same set of contrastive features organized in ostensibly the same sequence can constitute different structures in different prosodic positions within a language and across languages. So in English [tʃ] is an affricate when the fricative is performed as the release phase of the postalveolar stop (as in the phrase *catch it*), but it is a sequence of stop and fricative when the [ʃ] is phased later to allow the [t] to have its own alveolar target. This is a good example for talking about syllable boundaries, too, and leads them to look for the distributional cues that support the structural analysis of [ts] as an affricate in German.

With examples like this, one can emphasize that segments are not Platonic entities existing out there in nature separate from the phonologies of actual speakers of a language, and that the students cannot go about finding out whether [tʃ] is an affricate or a cluster in the same way that they might go about finding out whether initial voiceless stops in Indian English are aspirated or not. At this point they will be ready to be told the same fact about syllables, stresses, phrases, and so on. If there are native speakers of both English and Japanese in the class, for example, one can demonstrate that [ski] is a single syllable for the English speakers, but definitely two for the Japanese speakers. This then can lead into difficult cases in English, where reduction makes the syllable count difficult. For example, we can get them to design experiments to see how much a native speaker can reduce the first vowel in *supports* be before the word comes to be indistinguishable from *sports*. This in turn can illuminate one's definition of the lowest level of stress contrast in English.

In teaching about stress, I find it also helpful to give examples of languages where this most basic level of prominence contrast is marked by rather different phonotactic constraints. For example, since the language of class instruction is usually English for me, I can tell them about Swedish, where stress can be defined at the lowest level in terms of phonemic length contrasts (a stressed syllable has a phonemically long vowel or is closed by a long consonant), and about Mandarin Chinese, where it can be defined in terms of the tonal specification (i.e. an unstressed syllable is one with the so-called "neutral tone"). Then I can show how the seemingly very different phonological hallmarks for English and Mandarin are associated with nearly identical facts about phonetic coordination and undershoot, with unaspirated stops being voiced foot internally in both languages, and unstressed vowels often being reduced to a syllabic realization of a neighboring consonant, and so on. Also, this lets me make the important point that in English, as in Swedish, pitch accents can be anchored only to syllables that are stressed at this more basic segmentally defined level, and that pitch accent then defines another level of stress contrast above the basic level.

If the class includes many students who are taking phonetics as part of a program in speech pathology, another way to get across some of these points is to assign them to transcribe the productions of the very young children whom they see in the speech clinic. Since this is a standard part of the diagnosis for phonological disorder, they will see the relevance of this exercise immediately. The productions of very young children are often very difficult to transcribe because they have not yet achieved the precision in temporal coordination that supports a segmental analysis for adult speech. Pointing out this source of difficulty makes it easier to highlight the link between segmentation at all levels of the prosodic hierarchy and timing control.

Once the students have more background in phonetics and phonology, one can also have them exercise their analytic imaginations in a similar way by reading about nonsegmental analyses of consonant and vowel contrasts, such as Browman and Goldstein's Articulatory Phonology (e.g., [16]) and the Neo-Firthians (e.g. [17]) and the arguments about the segment in Steriade's recent work [18].

Above all, I think we must be humble about our current state of knowledge about the phonetics of prosody. The most helpful answer often will be simply: "We don't know, yet." But let us be sure to phrase it that way, to make the "we" inclusive, inviting the students into this grand enterprise.

REFERENCES

[1] Beckman, M. E. (1995). "Local shapes and global trends," This conference.

[2] Pierrehumbert, J. & Hirschberg, J. (1990). "The meaning of intonation contours in the interpretation of discourse," In P. R. Cohen, J. Morgan, & M. E. Pollack (Eds.), *Intentions in Communication*, pp. 271-311. Cambridge, MA: MIT Press.

[3] Cruttendon, A. (1974). "An experiment involving comprehension of intonation in children from 7 to 10," *Journal of Child Language*, vol. 1, pp. 221-232.

[4] Cruttendon, A. (1985). "Intonation comprehension in ten-year-olds," *Journal of Child Language*, vol. 12, pp. 643-661.

[5] Read, C., Zhang, Y. Nie, H., & Ding, B. (1986). "The ability to manipulate speech sounds depends on knowing alphabetic spelling," *Cognition*, vol. 24, pp. 31-34.

[6] Share, D., Jorm, A., Maclean, M., & Mathews, R. (1984). "Sources of individual differences in reading acquisition," *Journal of Educational Philosophy*, vol. 76, pp. 1309-1324..

[7] Bruce, G. (1977). *Swedish Word Accents in Sentence Perspective*. Lund: Gleerup.

[8] Kori, S. (1987). "The tonal behavior of Osaka Japanese: an interim report," *Ohio State University Working Papers in Linguistics*, no. 36, pp. 31-58.

[9] Chao, Y.-R. (1930). "A system of tone letters," *Le Maître Phonétique*, No. 30, pp. 24-27.

[10] Maddieson, I. (1990). "The transcription of tone in the IPA," *Journal of the International Phonetic Association*, vol. 20, pp. 28-32.

[11] Ladefoged, P. (1993). *A Course in Phonetics*. 3rd ed. London: Harcourt Brace Jovanovich.

[12] Pierrehumbert, J. B., & Beckman, M. E. (1988). *Japanese Tone Structure*. Cambridge, MA: MIT Press.

[13] Pitrelli, J., Beckman, M. E., Hirschberg, J. (1994). "Evaluation of prosodic transcription labeling in the ToBI framework," *ICSLP'94*, pp. 123-126.

[14] Black, A., & Campbell, N. (1995). "Predicting the intonation of discourse segments from examples in dialogue speech," ESCA Tutorial and Research Workshop on Spoken Dialogue Systems, Aalborg University.

[15] Jun, S-A. (1994). UCLA seminar.

[16] Browman, C., & Goldstein, L. (1989). "Articulatory gestures as phonological units," *Phonology*, 201-251.

[17] Local, J. K. (1992). "Modelling assimilation in non-segmental rule-free synthesis," In G. J. Docherty & D. R. Ladd, eds., *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*, pp. 190-223. Cambridge, UK: Cambridge University Press.

[18] Steriade, D. (1994). "On representing segmenthood," Paper presented at FLSM 4, University of Illinois, Urbana, IL.