Methods in Empirical Prosody Research

Edited by
Stefan Sudhoff, Denisa Lenertová, Roland Meyer,
Sandra Pappert, Petra Augurzky, Ina Mleinek,
Nicole Richter, Johannes Schließer

Offprint

Walter de Gruyter · Berlin · New York

Carlos Gussenhoven (Nijmegen, London)

Experimental Approaches to Establishing Discreteness of Intonational Contrasts

1 Introduction

In mainstream views of phonetics and phonology, the explanation of phonetic forms is divided into two parts. A phonetic form is first of all attributed to a cognitive representation, consisting of a grammatically legitimate combination of phonological elements. This part would remain constant if a speaker were to repeat the utterance, and is inherent to the particular linguistic expression pronounced by the speaker.

The variation that would be observable if a speaker were to repeat the pronunciation of the expression under a variety of circumstances is attributed to the phonetic implementation. Within this hypothetical set of repetitions, an infinite set of physical manifestations, there is no discreteness, only gradience. If we are to be able to state the phonological inventory and formulate the grammar of a language, the problem of distinguishing between phonological and phonetic differences must be solved.

Accordingly, a central issue in the analysis of intonation systems is that, given two utterances, how we are to know whether they came from the same phonological representation or from different phonological representations. The problem of distinguishing between these two sources of variation is particularly acute in intonation, because more so than variation in the segmental domain, phonetic variation in pitch is used for communication, and signals 'paralinguistic meaning' (Ladd, 1996).

The complicating factor here is that intonational meaning and paralinguistic meaning are in many ways similar. For instance, (1) and (2) have, on all accounts, different phonological structures, (1) having a high-pitched *se-* and (2) a mid-pitched or low-pitched one. In (1), the word *second* seems particularly important, and the example may sound as if there was a choice between *first* and *second*, while (2) might be a confirmation of the date of a meeting. The point is that (3) could be argued to have a different meaning from (1), in that it may be interpreted to mean that a third or fourth may also qualify ('it's minimally the second'), yet most analyses of English intonation would regard it as having the same structure as (1). Examples (1) and (3) would each consist of two accent peaks, analysed as

%L H*L H*L L% in Gussenhoven (2004) or as L+H* L+H* L-L% in Pierrehumbert (1980) and Beckman and Pierrehumbert (1986). The difference between (1) and (3) is that the second peak is carried out with greater pitch excursion in (3) than in (1). However, (2) is analysed as having a downstepped second peak, which is taken to be discretely different from the non-downstepped peak in (1) and (3).

(1) It's DEFinitely the SECond
%L H*L H*L L%

(2) It's DEFinitely the SECond
%L H*L !H*L L%

(3) It's DEFinitely the SECond
%L H*L H*L L%

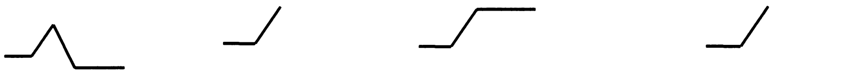
Ideally, if these analyses are correct, any contour that is broadly similar to those in (1) and (2) ought to be assignable by native speakers to either of these types, but contours that fall between those in (1) and (3) might create the impression of really being the same, and only assignable to either (1) or (3) on the basis of attentive listening to the pitch of the second peak.

2 Reviving an older experimental approach

In the large majority of cases, the structural discreteness that is assumed in analyses of intonation systems is rooted in native speaker intuition. It is only in the more subtle cases, such as when a language appears to have two kinds of rises or two kinds of falls, that the issue becomes problematic. In fact, the earliest operational criterion that was proposed is based on the explicit engagement of native speaker intuition. The IPO school of intonation analysis ('t Hart, Collier, and Cohen, 1990) distinguished between two levels of abstraction of the phonetic contour, referred to as 'close-copy stylization' and 'standard stylization'. In close-copy stylization, the contour is smoothed out with the help of the smallest number of straight-line sections as are needed to make the original and the stylization sound exactly the same. That is, there are no audible pitch differences, although the complete smoothing out of all microprosody may have an effect on the quality of the segments; for instance, raised pitch in a zone of 10 ms immediately before or after a voiceless consonant will make that consonant more clearly voiceless than without it (Kohler, 1990). There may be more than one solution, and the process is one of trial and error, but once the original contour and the close copy sound as if they have exactly the same pitch contour, they have 'perceptual equality'. An experiment with British English contours showed that this equality can be taken literally, in the sense that listeners hear no difference between the one and the other (de Pijper, 1983).

Standard stylization can be achieved once the straight-line sections obtained from the close copies have been sorted and averaged. In this way, a set of standardized movements can be formed. As the authors point out, "[t]he generalization should not go so far that it would do away with possible categorical differences between types of pitch movement" ('t Hart, Collier, and Cohen, 1990, p. 48). A contour can now be subjected to standard stylization by re-building it with the help of these standard straight line sections. In their grammar of Dutch intonation, there were five types of rises and five types of falls, each coming with a specification of whether they were accent-lending, where in the syllable their beginning should be aligned, and whether the slope is fixed or variable. The standard stylization and the close copy stylization should be 'perceptually equivalent', but will sound audibly different, because there may be substantial differences in pitch height and smaller differences in slope or timing between comparable sections of the contours: "For instance, some of the excursions may be audibly larger or smaller in the imitation than in the example, without an entire loss of similarity" ('t Hart, Collier, and Cohen, 1990, p. 47). An experimental operationalization of 'perceptual equivalence' was considered less straightforward than that of 'perceptual equality', but the authors were clear about what it might involve: the two audibly different contours, the close copy and the standard copy, must be heard as 'passable imitations' of each other by native speakers of the language. Because the authors felt that subjects might vary in their interpretations of when a contour was an imitation of some other contour, they never really explored the possibilities of a 'passable imitation' task and instead resorted to acceptability judgements, which at least could make sure that their standardized contours were realistic. The relevance to the concern in this chapter should be clear: in our terms, contours whose standard stylizations are the same have the same phonological representations, while contours with different standard stylizations are phonologically distinct.

Only recently has there been an experiment that used the criterion for 'perceptual equivalence' in an experimental task, i.e. one in which subjects had to judge whether one contour was a passable imitation of the other, Odé (2005). The issue here was whether the difference between two Russian intonation contours, which are readily heard as distinct when the accented syllable is non-final in the Intonational Phrase, are still distinct when the accented syllable is final, where truncation would appear to obliterate the difference between them. Impressionistic stylizations of the contours, provisionally analyzed as LH*L L% and LH* %, are given in Figure 1, from Odé (2005). The '%' by itself indicates an Intonational Phrase boundary without a boundary tone. The phonetic similarity between the two final-accented contours is high, as shown in Figure 2. The answer to the question whether these contours are still discretely different is therefore anything but trivial.



býlo véselo? 'was (it) fun?'	*býlo tepló (?)* 'was (it) warm(?)'	*býlo véselo (i interésno.)* '(it) was fun (and nice)'	*býlo tepló (i priiátno.)* '(it) was warm (and pleasant)'
LH*L L%		LH* %	

Figure 1: Stylized pitch contours of types LH*L L% (left pair) and LH* % (right pair) realized in utterance-final words on antepenultimate and ultimate syllables (from Odé, 2005)

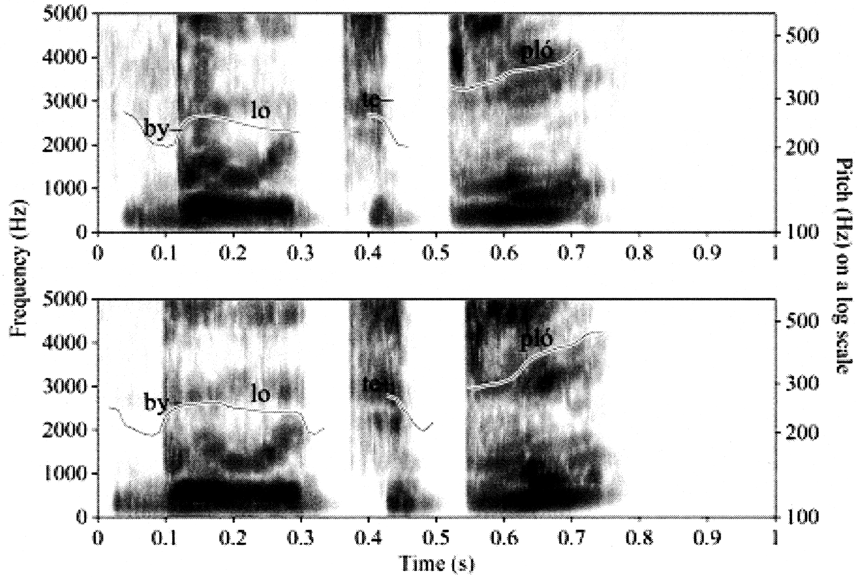


Figure 2: Spectrogram and pitch contour on a logarithmic scale of the utterance *býlo tepló* '(it) was warm' realized by a female speaker with contours LH* % (top) and LH*L L% (bottom) (from Odé, 2005)

Subjects were presented with naturally spoken utterances with the two intonation patterns as used in final and non-final position in the Intonational Phrase (IP). These had been obtained from eight speakers who had been instructed to read a list of sentences that had both final and non-final accents, once with LH*L L%

and once with LH* %. Pairs of stimuli were formed by combining utterances spoken by different speakers with contours from the same intonation contour (same-set pairs) and from different intonation contours (different-set pairs). The members of each pair were always readings of the same written sentences, and thus had the last pitch accent in identical locations. The results are summarized in Figure 3. Subjects' responses to contours taken from within an intonational set were judged to be passable imitations of each other in 72% of cases in both final and non-final positions. This may be looked upon as a baseline value for the recognition of phonological identity of phonetically different contours. Pairs of contours taken from different intonational sets were judged to be passable imitations of each other in only 7% of cases when the accent occurred in IP-internal position, where discrimination was expected to be high. This may be seen as a baseline value for easily distinguishable differences between phonological contours. Crucially, in final position, listeners judged the contours to be passable imitations of each other in significantly fewer cases than they did for contours taken from the same set, as shown in Figure 3. Clearly, if there were neutralization, the contours should have been judged the same in many more instances. The fact that the 'sameness' judgements did not reach the low baseline value of 7% is attributable to the very similar phonetic contours with which they are realized. Odé's conclusion therefore is that there is no neutralization, and that the truncation graphically represented in Figure 1 occurs in the phonetic implementation, and does not amount to the deletion of final L-tones, such as would cause LH*L L% and LH* % to become identical.

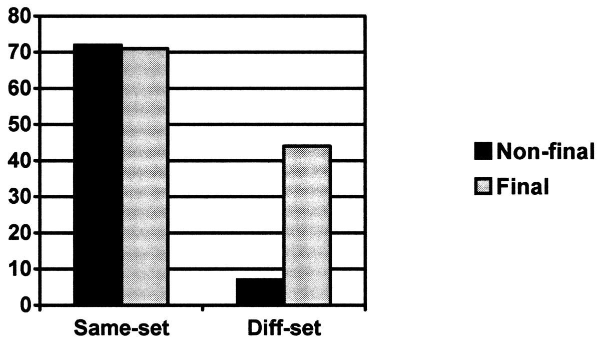


Figure 3: 'Passable-imitation' scores for pairs of utterances with LH*L L% contours or LH* % contours (same set) and pairs of utterances contrasting LH*L L% and LH* % contours (different set) in final and non-final position (after Odé, 2005)

The successful resolution of an issue that is crucial to the correct analysis of the intonation of Russian by means of this relatively straightforward procedure might profitably be explored further.

3 Implications for other approaches

In Gussenhoven (1999), I reviewed four experimental approaches which have sought to determine the nature of intonational contrasts on the basis of the performance of subjects in other experimental tasks. I briefly return to these below. They are 'categorical perception', the 'semantic difference task', the 'semantic scale task' and the 'imitation task'.

3.1 Categorical perception

If listeners interpret stimuli covering the phonetic continuum between two phonological categories as belonging either to the one category or to the other, and do not hear much difference between the stimuli that are interpreted as representing the same category, they perceive the contrast as 'categorical' (Liberman et al., 1957; Liberman et al., 1967). The classic case is that of the continuum of Voice Onset Times covering the canonical realisations of two stop phonemes, like /p/-/b/. Two separate tests are required to establish categorical perception. First, an identification task, in which listeners are presented with stimuli in a random order and are asked to assign each of them to either of the two phoneme categories, should reveal that there is a fairly abrupt perceptual shift somewhere along the continuum from one phoneme category to the other. Second, a discrimination task, in which listeners are presented with pairs of stimuli that differ by one acoustic step and are asked to indicate whether the two stimuli are the same or different, should reveal that between adjacent stimuli that are perceived as representing a single phoneme category listeners hear fewer distinctions than between the stimuli that lie on different sides of the perceptual boundary, even when acoustic differences are comparable. This ideal result is depicted diagrammatically in Figure 4, which was taken from Newport (1982).

As Newport points out, categorical perception is neither a necessary nor an exclusive property of linguistic contrasts. Phonological contrasts that are perceived categorically are voicing distinctions in plosives, while contrasts that are not may include vowel duration distinctions. Conversely, the non-linguistic distinction between a plucked string and a bowed string has been found to be perceived by listeners in the same way as the linguistic /p/-/b/ contrast (Cutting, Rosner, and Foard, 1976). In other words, categorical perception is not a property of linguistic stimuli, but of a class of acoustic stimuli that cross-cuts the linguistic-nonlinguistic distinction. While there have been a number of attempts to demonstrate the usefulness of this task to discovering intonational contrasts, these do not have the same power of conviction as Odé's result reviewed in Section 2. It may thus be more effective to appeal to listener's intuitions in a more direct way, as in the 'passable imitation' task, than to probe the listener's linguistic knowledge through a test that emphasizes his ability to discriminate between phonetically different forms, since the latter procedure may bypass the distinction the test is aimed to reveal, that between non-structural and structural phonetic differences.

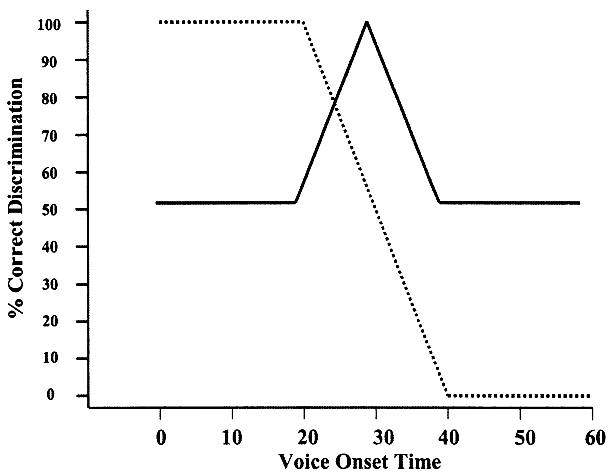


Figure 4: Idealized discrimination function in relation to the idealised identification functions for stimuli varying in Voice Onset Time, from Newport (1982)

3.2 Semantic difference task

Semantic judgements can be gradient, for which a semantic scale is used running from the full presence of a meaning ('friendly') to its opposite ('not friendly'), or categorical, in which case subjects have to choose from a set of two or more meanings. (On rating scales, see Rietveld and Chen, this volume.) The 'semantic difference task' amounts to asking subjects to choose one meaning from a set, typically a set of two, of stimuli that differ in their f0 contour only. An *a priori* difficulty with this task is that, as explained above, phonetic differences in intonation are typically meaningful. Ladd and Morton (1997) ran a series of experiments to test the hypothesis that English has a discrete difference between H and an 'extra High', a difference that is discrete in some tone languages. The difference between a high and a low pitch peak can be associated with different meanings in English, as shown in (4a) and (4b).

(4) a. A: He's Armenian.
B: He's iRANian. (with high H*+L peak: contradiction)

b. A: His name is Kameiny.
B: He's iRANian. (with neutral H*+L peak: neutral statement)

The subjects in this experiment had no difficulty associating higher peaks with a contrastive interpretation and lower peaks with a neutral interpretation. As made clear by Ladd and Morton, however, the fact that listeners assign different meanings to phonetically different forms need not mean that the two forms are discretely different, even if they are judging randomly presented sets of stimuli that represent a phonetic continuum between two forms. Such a result is a probable outcome of any task with two response categories (with or without a 'don't-know' category). The mere fact that the curve plotting the pooled responses for each of the two meaning has an S-shape, indicating that each end of the continuum is associated with each of the meanings, therefore tells us little about the nature of the contrast.

Before Ladd and Morton (1997), it was not uncommon to look upon the establishment of a semantic contrast as sufficient proof of a structural difference. A case can perhaps be made for meaning contrasts that are evidently non-paralinguistic. However, this line of argument is not without problems. First, paralinguistic meanings have not so far been clearly defined. They are usually called 'paralinguistic' because they are expressed in the phonetic implementation rather than the linguistic structure, and using paralinguistic meaning as criterial therefore risks circularity. But even the establishment of non-attitudinal, non-affective meaning differences will not guarantee that the forms involved are categorically different. Given the right context, major differences of interpretation can be cued through pitch range differences that are phonetic in nature. For instance, the past tense of *think* in *I THOUGHT so* can be interpreted as 'I already thought so at the time, and I still think so', or as 'I thought so at the time, but not any longer'. In the first case, the speaker is proved right, in the second case he is proved wrong, as shown in (5a) and (5b), respectively.

(5) a. A: Did you see a shooting star?
B: I THOUGHT so. And I was right.
b. A: Did you see a shooting star?
B: I THOUGHT so. But I was wrong.

It is imaginable that if we had a series of stimuli varying in peak height on the word *thought*, the lower peaks would be judged to be appropriate for meaning (5a) while the higher peaks would be associated with meaning (5b), in a forced-choice task. That is, even though the difference in form is to be accounted for in the phonetic implementation, pragmatic inferencing may lead to a difference of interpretation that involves the presence or absence of a shooting star.

Nash and Mulac (1980) carried out exactly this experiment, with the addition of a 'don't-know' response category, except that they didn't vary peak height, but peak alignment (as well as pitch height on *I*, which variation didn't produce any interpretable results). The peak alignment difference correlated with the meaning difference in (5a,b), such that early peaks were associated with (5a) and late peaks with (5b). In their case, the difference in form was in fact phonological. Contour (6b) is a more complex form of (6a), due to the addition of the prefix-L to the fall of (6a), adding an element of 'significance' to it (Gussenhoven, 2004, ch. 15). However, the same result might arguably have been obtained with stimuli that vary in peak height and which are phonologically identical.

(6) a.

[I THOUGHT so]
H*L H% ('fall')

b.

[I THOUGHT so]
L*HL H% ('rise')

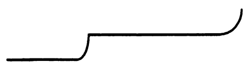
3.3 Semantic scaling task

When listeners are asked to judge the extent to which an utterance conveys a particular meaning, they give their estimate on some scale, either discrete ('7-point scale') or analogue (see Rietveld and Chen, this volume). Typically, these semantic scales are used in intonation research to determine paralinguistic meanings conveyed by pitch contours. However, Gussenhoven and Rietveld (2000) illustrate an investigation of a paralinguistic meaning that provided information about the tonal composition of two classes of pitch contour. The investigation started from the assumption that differences in pitch span affect H-tones and the accent-lending L*-tones differently. While higher realisations of H-tones correspond to increases in pitch range, L* may be lower as pitch range increases (Liberman and Pierrehumbert, 1984). Gussenhoven and Rietveld (2000) exploited this fact to address the nature of the difference between the 'low rise' and the 'high rise' in Dutch. The 'low rise' begins low, rises to mid at the end of the accented syllable or immediately after the accented syllable, and has a further rise on the last syllable, as shown in (7a). The 'high rise' begins mid, usually after a low unaccented stretch which may continue as far as the CV-boundary of the accented syllable, then continues at the same mid pitch until it reaches the final syllable, where a final rise to high occurs, as shown in (7b). When the accented syllable is IP-final, the rise due to H% will occur in the second half of the accented syllable.

(7) a.

[Zijn er meLOENen teveel]
L*H H%
'Are there too many melons?'

b.



[Zijn er meLOENen teveel]
H* H%

Using stimuli consisting of manipulated versions of four utterances, two with final accent and two with non-final accent, the authors asked subjects to rate the accented syllables for 'surprise'. The sentences used were lexically fairly neutral, and no attempt was made to choose sentences that expressed any form of surprise. Nine of the stimuli represented versions of the low rise, in which three values for the beginning of the rise (80, 90, and 100 Hz) had been crossed with three values for the end of the rise (185, 200, and 215 Hz), and nine stimuli represented versions of the high rise, in which three values of the beginning of the rise (130, 145, and 160 Hz) had been crossed with three values for the end of the rise (185, 200, and 215 Hz). (In the low rise contours, the mid level stretch was fixed at 145 Hz.)

It appeared that perceived surprise correlated first of all with the height of the end of the contour, i.e. with the pitch of H%. More importantly, it turned out that perceived surprise corresponded with *lower* beginnings of the low rise (represented by L*) and *higher* beginnings of the high rise (represented by H*). This finding supports the hypothesis that (7a) and (7b) represent a phonological contrast, and also that the pitch accents are represented by L* and H*, respectively. If the stimuli were to form a set of phonetically different pronunciations of the same phonological representation, perceived surprise would – incongruously – follow a U-shaped pattern, with a high degree of surprise for low and high beginnings and a low degree for mid-pitch beginnings. No such U-shaped behaviour was found for the end-point of the rise, which is linearly related to the level of perceived surprise and which is represented phonologically by a H%-tone in both contours. That is, perceived surprise increases with increases in pitch range, and pitch range increases are achieved by lowering L* and raising H-tones.

This test, while yielding clear results in the case described, has more limited applicability than the imitation task, and depends on there being a paralinguistic meaning that is sensitive to pitch range expansion, like 'surprise', as well as a research question that involves a difference between H* and L*. In fact, other meanings that were elicited with the same stimuli, like 'emphatic' and 'insistent', didn't yield the same results. Also, it is not clear that all L-tones go down when the pitch range is expanded, or even if L*-tones go down in all languages.

3.4 Imitation task

A new experimental approach to establishing discreteness was adopted by Pierrehumbert and Steele (1989). They were concerned with the same contrast that was investigated by Nash and Mulac (1980), shown in (8) with ToBI transcription (cf. Beckman and Pierrehumbert, 1986; for discussion of the meaning of (8b) see also Hirschberg and Ward, 1992).

(8) a.

[Only a MILLionaire]
H*+L L-H%

b.

[Only a MILLionaire]
L*+H L-H%

The peak of (8a) occurs inside the accented syllable, while that in (8b) will typically occur during the following unstressed syllable. The authors produced a continuum of 15 peak alignments in a set of resynthesised stimuli ranging between (8a) and (8b) in steps of 20 ms. Subjects were asked to imitate each of a series stimuli presented in a randomized order, paying particular attention to the intonation pattern. Pierrehumbert and Steele argued that if subjects were capable of reproducing the continuum in their imitations, the difference must be gradient. However, if subjects were to produce a binomial distribution of the peak times in their imitations, then the difference must be categorial. Four subjects produced 30 imitations of each stimulus, giving a total of 1800 utterances. For each utterance, the peak time was measured relative to the release of [m]. The results supported a categorical contrast: the distribution of the peak timings tended to be binomial. Because these subjects did not produce a continuum of peak alignments, but tended to have many early and late peak alignments and fewer intermediate ones, a discrete interpretation of the contrast would appear to be called for. Since other research has shown that speakers can in fact produce gradually different f0 *peak heights* in production experiments in which they are asked to pronounce sentences at a prominence level as specified by a scale value on a 10-point scale (e.g. Liberman and Pierrehumbert, 1984), we can have some confidence that results obtained in Pierrehumbert and Steele's experiment in fact do suggest discreteness.

However, some qualification of the usefulness of the task is in order after the work by Redi (2003), Dilley (2005), who used four phonetic continua in English. Two of these were peak continua and two were valley continua, as listed in (9). The continua consisted of two slopes, up and down for the peaks, and down and up for the valleys, where the peak or the valley was shifted in 15 steps of 25 ms. The pitch before and after the peaks was low, and before and after the valleys high. The peak and the valley continua each appeared over a sequence of stronger-weaker as well as a sequence of weaker-stronger syllables. Continuum (9b) was intended to represent phonetic variation, while the other three were assumed to correspond to phonological contrasts.

(9)	a. *To Monrovia*	weak-strong peak continuum between a peak on *mon* and a peak on *RO*, interpretable as H on *mon* preceding either a downstepped !H* (early peak) or a non-downstepped H* (late peak) on *RO*
	b. *Too minglingly*	strong-weak peak continuum between a peak on *MING* and a peak on *ling*, both to be interpreted as H*, but with varying alignments of the peak
	c. *To Monrovia*	weak-strong valley continuum between early valley on *mon* and late valley on *RO*, interpretable as L+H* (early) or L* (late) on *RO*
	d. *They're non-linguistic*	strong-weak valley continuum between early valley on *NON* and late valley on *ling*, interpretable as L*+H on *non* (early) and L+H* on *GUIST* (late)

The results showed that subjects imitated peaks in a bimodal fashion, but that valleys were reproduced with a continuum of alignments of the low points. This suggests that the imitation task needs to be re-evaluated, since it is possible that the peak imitations are bimodal regardless of the nature of the input, but that valley imitations are not. Continuum (9b) was expected to be a gradient one, but perhaps because only the alignment of a single point was manipulated, there may have been ambiguous or unnatural stimuli. As suggested by Bob Ladd (Dilley, 2005, 103), (9b) may have been interpreted as representing H* in the early peaks and L*+H for the late peaks. Instead of manipulating a single point, the continuum might be constructed by 'morphing' one canonical contour to the other, while stopping at a number of points on the way. Additionally, it may be worth investigating whether a different task, one that more directly addresses speaker intuitions, gives better results. For instance, in their imitations, subjects could be asked to 'correct' the pronunciation of stimuli that are presented as having not quite acceptable intonation contours, rather than to reproduce the contours of the stimuli as best they can.

The imitation task was significantly expanded by Kochanski, Braun, Grabe, and Rosner (Ms.; cf. Kochanski, this volume), who designed an experiment in which subjects imitated their own imitations of an initial artificial contour. The first contour was deliberately constructed such that it did not obviously represent a well-formed English intonation contour. After four recursions, the end products represented a smaller number of well-formed contours, whereby each iteration approached the end product more closely. The task can reveal the non-phonological nature of phonetic differences if subjects' imitation of different contours end up as the same contour, and show that differences are discrete if they remain different in their end products.

4 Conclusion

A reconsideration of the methodology proposed by 't Hart, Collier, and Cohen (1990) to establish phonological identity of phonetically different intonation contours suggests that subjects' tasks in perception and imitation experiments seeking to establish dicreteness should be more explicitly formulated so as to tap into subjects' intuitions. The aim here is to reduce the effect of phonetic differences that fall within one phonological category on the scores. Judgements involving sameness or difference can be more focused on structural differences by asking whether one contour can serve as a passable imitation of the other, instead of asking whether they are different, and imitation tasks can ask the subject to correct a phonetic contour that lies between hypothetical categories to its 'intended' pronunciation instead of asking the subject to reproduce the stimulus as accurately as they can.

References

Beckman, M. and J. Pierrehumbert (1986): Intonational structure in English and Japanese. Phonology Yearbook 3, 255-310.

Cutting, J., B. Rosner, and C. Foard (1976): Categories and boundaries in speech and music. Quarterly Journal of Experimental Psychology 28, 361-378.

de Pijper, J. R. (1983): Modelling British English Intonation. Dordrecht: Foris.

Dilley, L. (2005): The phonetics and phonology of tonal systems. Ph.D. Dissertation, MIT.

Gussenhoven, C. (1999): Discreteness and gradience in intonational contrasts. Language and Speech 42, 281-305.

Gussenhoven, C. (2004): The Phonology of Tone and Intonation. Cambridge: Cambridge University Press.

Gussenhoven, C. and T. Rietveld (2000): The behavior of H* and L* under variations in pitch range in Dutch rising contours. Language and Speech 43, 183-203.

Hirschberg, J. and G. Ward (1992): The influence of pitch range, duration, amplitude and spectral features on the interpretation of the rise-fall-rise intonation contour in English. Journal of Phonetics 20, 241-251.

Kochanski, G., B. Braun, E. Grabe, and B. Rosner (Ms.): Probing the Mental Representation of Intonation.

Kohler, K. (1990): Macro and micro F0 in the synthesis of intonation. In: J. Kingston and M. E. Beckman (eds.): Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech. Cambridge: Cambridge University Press, 115-138.

Ladd, D.R. (1996): Intonational Phonology. Cambridge: Cambridge University Press.

Ladd, D.R. and R. Morton (1997): The perception of intonational emphasis: Continuous or categorical? Journal of Phonetics 25, 313-342.

Liberman, A.M., K.S. Harris, H.S. Hoffman, and B.C. Griffith (1957): The discrimination of speech sounds within and across phoneme boundaries. Journal of Experimental Psychology 61, 379-388.

Liberman, A.M., F.S. Cooper, D.P. Shankweiler, and M. Studdert-Kennedy (1967): Perception of the speech code. Psychological Review 74, 431-461.

Liberman, M.Y and J.B. Pierrehumbert (1984): Intonational invariance under changes in pitch range and length. In: M. Aronoff and R.T. Oehrle (eds.): Language sound structure: Studies in phonology presented to Morris Halle. Cambridge, MA: MIT Press, 157-233

Nash, R. and A. Mulac (1980): The intonation of verifiability. In: L.R. Waugh and C.H. van Schooneveld (eds.): The melody of language: Intonation and prosody. Baltimore: University Park Press, 219-241.

Newport, E.L. (1982): Task specificity in language learning? Evidence from speech perception and American Sign Language. In: E. Wanner and L.R. Gleitman (eds.): Language acquisition: The state of the art. Cambridge: Cambridge University Press, 450-486.

Odé, C. (2005): Neutralization or truncation? The perception of two Russian pitch accents on utterance-final syllables. Speech Communication 47, 71-79.

Pierrehumbert, J. (1980): The phonology and phonetics of English intonation. PhD dissertation, MIT.

Pierrehumbert, J.B. and S.A. Steele (1989): Categories of tonal alignment in English. Phonetica 46, 181-196.

Redi, L. (2003): Categorical effects in the production of pitch contours in English. In: Proceedings of the 15th International Congress of the Phonetic Sciences, Barcelona, 2921-2924.

't Hart, J., R. Collier, and A. Cohen (1990): Perceptual Study of Intonation: An Experimental-Phonetic Approach to Speech Melody. Cambridge: Cambridge University Press.