Linguistica Silesiana 38, 2017
ISSN 0208-4228

ARKADIUSZ ROJCZYK
arkadiusz.rojczyk@us.edu.pl
ANDRZEJ PORZUCZEK
andrzej.porzuczek@us.edu.pl
University of Silesia

DISCRIMINATION OF ENGLISH TONE CONTOURS BY POLISH LEARNERS

In the current study, we test the discrimination of four basic English tone contours in monosyllabic words by Polish learners using an AXB task and we compare these results to the results of an identification test. Discrimination does not require access to phonological labels and is claimed to tap core auditory mechanism. Relatively high discrimination performance by Polish learners and poor identification performance indicate that difficulties with correct identification of English tones are solely difficulties with labelling.

1. Introduction

Intonation is one of these features of speech that are regarded as important in communication but on the other hand they are extremely difficult to teach and learn. For instance, Dalton and Seidlhofer (1994: 73) claim that intonation patterns are "particularly important in discourse" while "at the same time they are particularly difficult to teach". Likewise, according to Grabe *et al.* (2005: 311), intonation "has been described as the most difficult aspect of a foreign language to acquire and is held responsible for numerous instances of miscommunication between native and non-native speakers." Consequently, it has been handled from the pedagogical perspective by authors of publications for English learners, e.g. O'Connor and Arnold (1973), Cook (1968), Brazil *et al.* (1980), Brazil (1994, 1997), Bradford (1988), Dalton and Seidlhofer (1994), Underhill (1994).

On the other hand, intonation is not very willingly included in actual teaching. This reservation is expressed in the following quotations, which indicate that intonation:

– “is widely regarded as slippery” (Ranalli 2002),
– “is one of those territories where many language teachers fear to tread” (Setter 2005),
– “is not teachable and possibly not learnable either” (Taylor 1993).

Several possible reasons for this attitude among teachers deserve mention. The most obvious ones are the complexity and variability of intonation patterns (cf. Roach 2000, Grabe *et al.* 2005) and the difficulty, even for native speakers, of conscious analysis and planning of intonation patterns, despite usually successful subconscious processing of familiar patterns (cf. Bradford 1988, Brazil 1994, Kelly 2000). Moreover, according to Cauldwell and Allen (1997: 2), “people vary in their ability to hear intonation patterns, and there are quite often disagreements between trained listeners about what they hear in a speech sample”.

As a result of the large variation among native speakers in using intonation, textbook authors find it difficult to propose a ready-made list of ‘correct’ patterns to be taught. Even if this problem is finally resolved, learners need the cognitive ability to perceive, identify and, eventually, produce the contours according to the decisions they have made. The knowledge of target language relations between particular tones and their pragmatic meanings is only a necessary but not a sufficient condition of successful communication in a foreign language (FL). Communicative failure that can be attributed to intonation may result from the following problems at successive stages of conversation:

– ignorance of FL patterns resulting in using native language (L1) intonation (perception and production)
– wrong contour identification (perception)
– wrong pragmatic interpretation of a contour (perception)
– wrong contour choices (pre-production stage)
– inability to produce a desired pattern (production)

Most studies of intonation investigate the production of native speakers or second language (L2) learners but there is less interest in the crucial ability to discriminate different tones by ear. One reason to neglect that field of study is the easily observed fact that language users are usually sensitive to intonation cues, they can interpret them correctly in L1, and they can fairly well imitate the patterns they hear. However, as observed by Kelly (2000: 86), intonation is “an aspect of language we are very sensitive to, but mostly at an unconscious level,” whereas learning a foreign language in classroom setting requires conscious use and practice of language patterns.

Finally, at any stage of education, we must allow for interference, the influence of L1 on the process of foreign language learning. A large number of studies have demonstrated how a native language can shape the perception of a foreign language. Situations where a language-specific phonology models the perception of non-native contrasts have been documented for both segments and prosody. For example, Grabe et al. (2003) found the effect of native language experience on the perception of tone contours in British English by Spanish and Chinese listeners. The results showed that, despite similarities in the performance of the tested groups, there were some language-specific differences in the perceptual organization. It was speculated that tones are generally processed by a universal auditory mechanism, but native language experience can modify the perceptual output.

Language prosody, including FL intonation, forms a part of practical phonetics course of university English studies. Polish learners of English have repeatedly been observed to have difficulties with identifying English tone contours, especially those with changing trajectory, such as fall-rise and rise-fall. It is an open question whether those difficulties result from the inability to correctly identify and label the taught contours, or whether the sources of this inability are located deeper in the psychoacoustic processing of prosody. If the problem is only on the labelling level, then discrimination of such contours should be high.

In our study we focus on the perception of basic intonation patterns by advanced Polish learners of English. In particular, we investigate how efficient they are at recognizing the patterns, trying to track possible influence of prior practical English prosody training and parallel tone language (Mandarin) learning.

2. Stimuli

We tested the recognition of four basic intonation contours: rise, fall, rise-fall, and fall-rise (Brazil *et al.* 1980, Roach 2000: 155-158). Four monosyllabic English words (*yes*, *no*, *good*, *bad*) were used as tone carriers, which yielded 16 word-tone combinations. They were recorded by a qualified phonetician in the Acoustic-Phonetic Laboratory at the University of Silesia. The signal was captured at 44100 Hz (24 bit quantization) through a headset dynamic Sennheiser HMD 26 microphone fed by a USBPre2 (Sound Devices) amplifier. The pitch range in the stimuli was approximately 5-6 semitones (a fourth) for simple tones (rise and fall) and 12 semitones (an octave) for the complex ones. The stimuli we normalized for intensity at 70dB for the experiment. Figure 1 shows the four tone patterns in the four test words.

3. Participants

The participants were 67 English studies majors at the University of Silesia, 28 males and 39 females, aged 19-27 ($M = 20.9$). Twenty-six second-year students had completed a practical English phonetics course including intonation practice, while the remaining fourty-one first-year students had not been provided any training. Thirty-six participants studied Chinese as the second foreign language, the other thirty-one participants studied a different second foreign

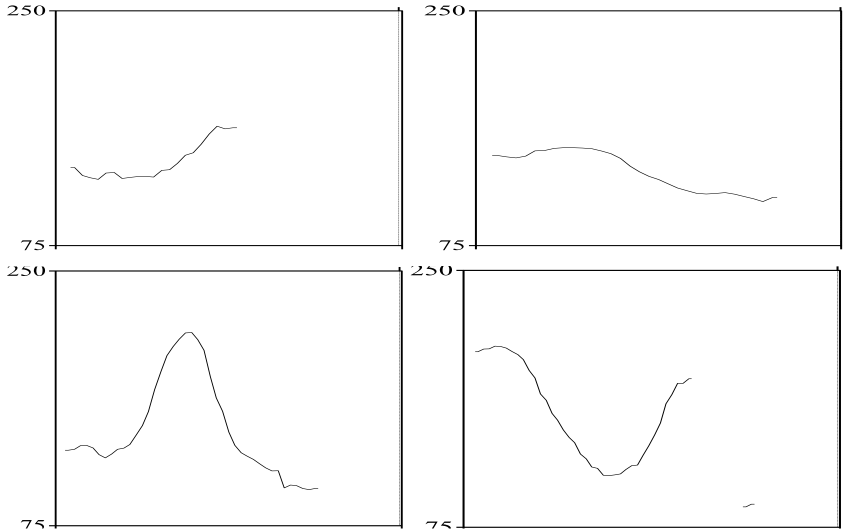

Figure 1. Examples of intonation patterns used for the experiment.
Rise in *yes* (top left), fall in *no* (top right), rise-fall in *good* (bottom left), fall-rise in *bad* (bottom right)

language. Such a participant design allowed direct comparisons (1) between listeners with prior training in tone recognition and those without training; (2) between listeners who had experience with a tone language and those that had no such experience. None of the participants reported any hearing or speech disorders. All of them had normal or corrected-to-normal vision.

4. Procedure

The experiment consisted of two perception tasks: identification and AXB discrimination. Both tasks were designed and run in Praat (Boersma 2001), using a 17-inch monitor with the audio signal fed by Philips SBC HP840 headphones at a comfortable listening level. In the identification task, the listeners were to select one of the intonation patterns displayed on the screen that would represent the pattern they had just heard. Graphic representations, together with the names of tone contours, were used to help the listeners recognize the options they were to choose from. Figure 2 shows the experimental panel in the identification task.

In each trial the listeners could use a ‘play again’ option once. Eight familiarization trials with stimuli not used in the analysis preceded the identification of target stimuli. There were 4 target words with 4 tone patterns, which altogether made 16 target stimuli for identification.

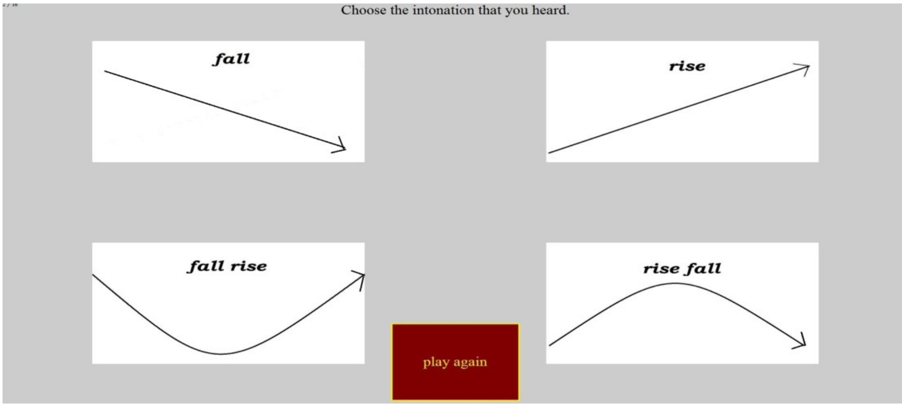

Figure 2. The experimental panel in the identification task

In the AXB discrimination task the listeners were required to indicate whether the second word had the same intonation pattern as the first or the third word by clicking an appropriate dialog box. The stimuli were arranged in all possible combinations, such as AAB, ABB, BBA, BAA. All triads had different words to avoid lexical biasing of responses, for example *good* (rise) – *yes* (fall) – *bad* (fall) or *no* (fall-rise) – *good* (fall-rise) – *yes* (rise-fall). The inter-stimulus-interval was 500 ms. As in the identification task, the listeners could resort to a 'play again' option once for each triad. Figure 3 shows the experimental panel in the discrimination task.

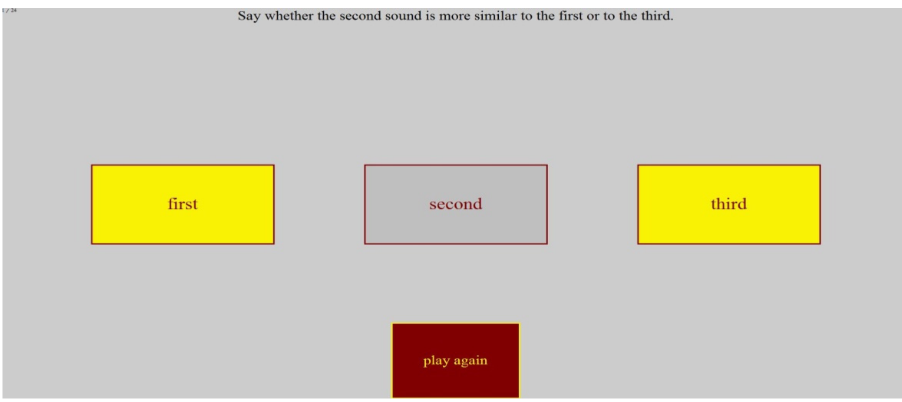

Figure 3. The experimental panel in the discrimination task

The order of tasks was fixed for each listener in that each session started with the identification task followed by the discrimination task. Each individual session lasted approximately 10 minutes.

5. Analysis and results

The results were analysed differently for the two tasks. For the identification task the proportion of correct identifications was calculated together with confusion matrices. For the discrimination task the sensitivity index d' (Green and Sweets 1966) was calculated. D' separates the distribution of sensitivity and bias in the listeners' discriminations. In other words, it takes into account not only the correct discriminations, but also looks into the incorrect responses (Macmillan and Creelman 1991). It is computed by subtracting the z-score that corresponds to the false alarm rate from the z-score that corresponds to the hit rate (Stanislaw and Todorov 1999).

5.1. Identification

The total proportion of identification accuracy for all tone contours was 54%.

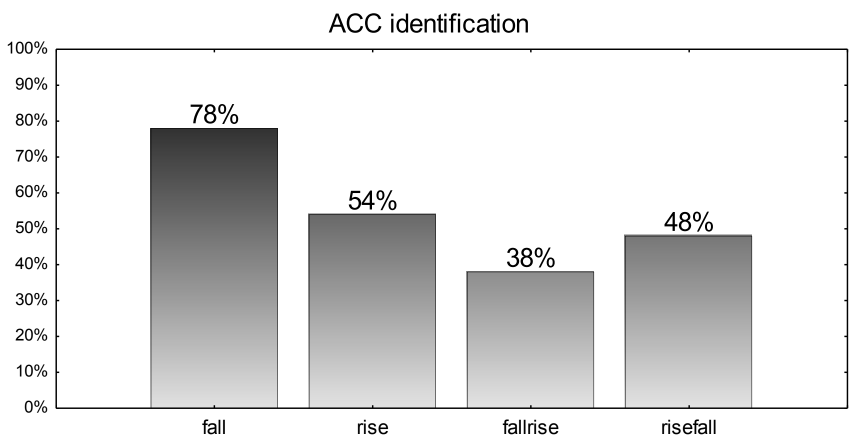

Figure 4. Identification accuracy for individual tone contours

The results clearly indicate that simple tone patterns were better recognized than complex tone patterns. The easiest to identify was 'fall' (78%) followed by 'rise (54%) and 'rise-fall' (48%). The most difficult to identify was 'fall-rise' (38%).

The confusion matrix in Table 1 shows the pattern of incorrect identifications for each individual tone contour.

The analysis of confusion patterns revealed that the listeners misperceived the test tones within the category of complexity. Simple tones, such as 'rise' and 'fall' were confused with each other much more frequently than with complex tones. The 'fall' tone was misidentified as 'rise' in 18% of the trials and the 'rise' tone was misidentified as 'fall' in 38% of the trials. Similarly, complex tones were largely confused with each other rather than with simple tones. The

Table 1. Confusion matrix for identification/misidentification of individual tone patterns

TONE	fall	rise	fall-rise	rise-fall
fall	78.0%	18.0%	0.7%	3.3%
rise	38.4%	53.7%	4.9%	0.3%
fallrise	2.6%	19.0%	38.4%	39.9%
risefall	5.6%	15.7%	31.0%	47.8%

‘fall-rise’ was incorrectly identified as ‘rise’fall’ in 40% of the tokens, which is even more frequently than it was correctly identified as ‘fall-rise’ (38%). The ‘rise-fall’ tone was confused with the ‘fall-rise’ in 31% of the trials.

For the between-group comparisons, the correct identifications were treated as a continuous variable ranging from minimum 0 correct identifications to maximum 16 correct identifications. A Kolmogorov-Smirnov test was used to test normality of distribution of the dependent variables. The distribution of overall identifications was normal [$D(67) = .11$, $p > .05$]. Between-group tests also indicated that the distributions were normal in the first (no phonetic training) [$D(41) = .17$, $p > .05$] and second (phonetic training) year of studies [$D(26) = .14$, $p > .05$] as well as in the group with Chinese as a second foreign language (Chinese group) [$D(36) = .14$, $p > .05$]] and in the group with a different foreign language (non-Chinese group) [$D(31) = .12$, $p > .05$]. Accordingly, the results were analysed using a factorial two-way ANOVA with phonetic training as one independent variable (first-year / second-year) and programme of studies (second foreign language) as another independent variable (Chinese / non-Chinese). There was a main effect of phonetic training on the correct identification [$F(1, 63) = 10.32$, $p < .01$, $\eta_p^2 = .14$], indicating that the second-year listeners performed significantly better ($M = 10.19$; $SE = 0.62$) than the first-year students ($M = 7.76$; $SE = 0.45$). When represented as accuracy percentage, the first-year students identified 48% of the tokens while the second-year students identified 64% of the tokens. There was no main effect of programme [$F(1, 63) = .20$, $p > .05$], showing that the listeners learning Chinese as a second language were not better at labelling the tested tones. In fact, the mean of correct identifications for the listeners learning Chinese was slightly lower ($M = 8.6$; $SE = 0.52$) than the one for the listeners with other second foreign languages ($M = 8.81$; $SE = 0.6$). The accuracy percentage was 54% for the students of Chinese and 55% for the students of other languages.

The interaction between phonetic training and programme was significant [$F(1, 63) = 5.73$, $p < .05$, $\eta_p^2 = .08$], revealing differences in tone identification between the listeners learning Chinese and those learning other foreign languages depending on whether or not they had received phonetic training.

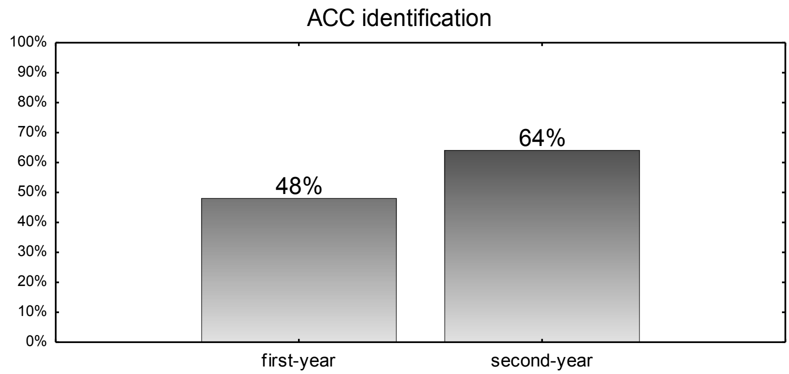

Figure 5. Identification accuracy for the phonetic training (first-year / second-year)

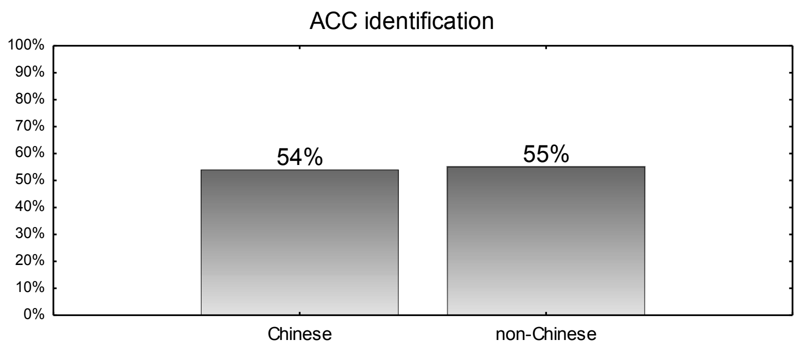

Figure 6. Identification accuracy for the second foreign language (Chinese / non-Chinese)

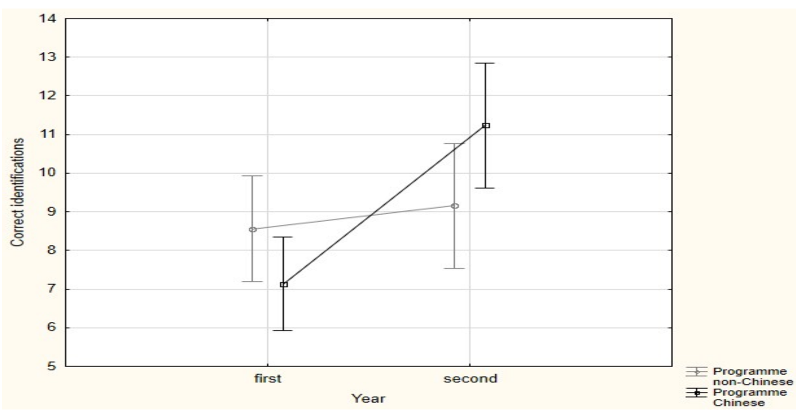

Figure 7. Interaction between phonetic training (first-year / second-year) and programme (Chinese / non-Chinese)

This interaction demonstrates that the listeners studying Chinese as a second foreign language and the listeners that studied other second foreign languages benefited differently from phonetic training. Post-hoc Bonferroni tests showed that the group studying Chinese in the second year performed significantly better in identification than the same group in the first year ($p < .001$). The same gain was not observed for the group studying other foreign languages ($p > .05$).

5.2. Discrimination

Overall results from the discrimination task show that the listeners discriminated contours at a near ceiling level of 92% with the mean sensitivity index d' of 2.67 (*SD* = .54). It suggests that discrimination was almost perfect. Unlike the results obtained in identification, there were no observable differences in the discrimination of individual tones. The discrimination accuracy ranged from 88% for fall-rise to 95% for rise-fall. Both simple and complex tone contours approached a ceiling level of discrimination.

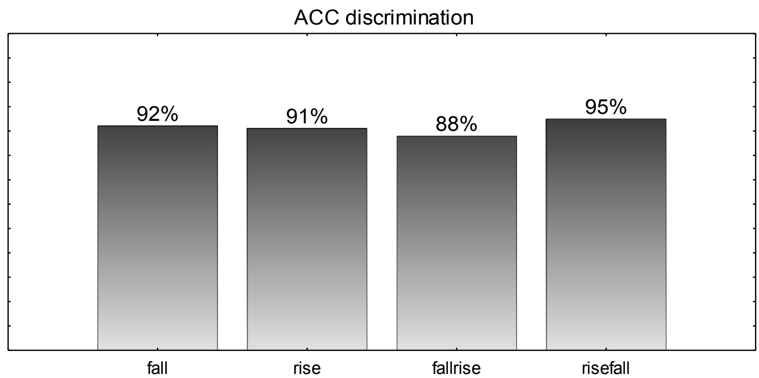

Figure 8. Discrimination accuracy for individual tone contours

For the between-group comparison the measure of sensitivity index d' was used. A Kolmogorov-Smirnov indicated that the data were not normally distributed [$D(67) = .21$, $p < .01$]. There was a negative skew of -.82 due to general high discrimination of the stimuli. As a result, all parametric tests were followed by non-parametric tests to verify the outputs that may be biased by non-normality of the distribution. For parametric tests a factorial two-way ANOVA was used with phonetic training (first-year / second-year) and second foreign language (Chinese / non-Chinese) as two independent variables. For non-parametric tests Mann-Whitney U tests will be reported together with medians. There was no main effect of phonetic training on discrimination of the tested tones [$F(1, 64) = 1.41$, $p > .05$; $U = 416.5$, $p > .05$], indicating that the listeners without phonetic training in the first year (*M* = 2.61; *SE* = .08; *Mdn* = 2.63) did not differ in their discriminations from the listeners after phonetic training in the second year (*M* = 2.77; *SE* = 0.11; *Mdn* = 3.08). There was no main effect of second foreign language [F(1, 64) = 1.07, p > .05; *U* = 469.5, p > .05], showing that learning Chinese did not affect the performance. The mean sensitivity index for the listeners with Chinese as a second foreign language was *M* = 2.73 (*SE* = 0.09; *Mdn* = 2.94),whereas for the listeners with other foreign languages it was *M* = 2.6 (*SE* = 0.1; *Mdn* = 2.63). Figures 9 and 10 show the accuracy percentage of discrimination for phonetic training and second foreign language.

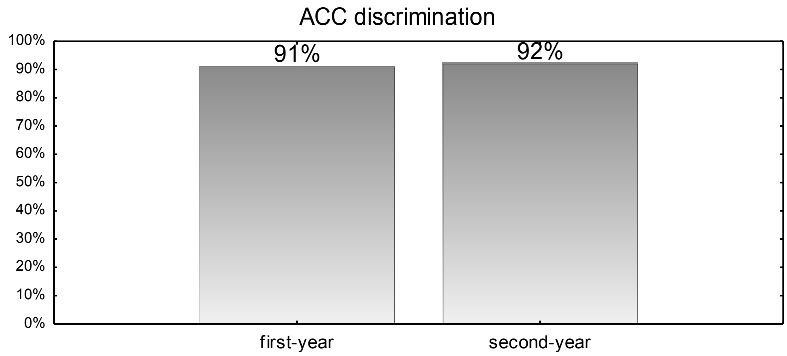

Figure 9. Discrimination accuracy for the phonetic training (first-year/second-year)

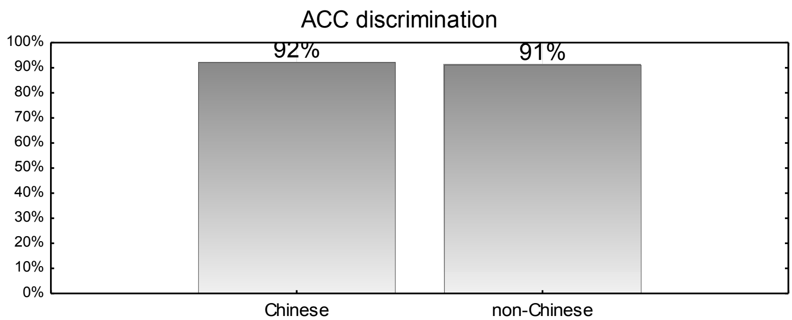

Figure 10. Discrimination accuracy for the second foreign language (Chinese/non-Chinese)

6. Discussion

The analysis of the results from the identification task shows that phonetic labelling of tone patterns posed serious difficulties for the listeners. The mean percentage of 54% indicates that the listeners performed above the chance level of 25%, however nearly 50% of the stimuli were misclassified. It reflects earlier observations from teaching prosody that very few learners manage to learn to identify tone contours at a satisfactory level. The analysis of the identifications for particular tone patterns reveals that simple tones (fall, rise) were identified much more accurately than complex tones (fall-rise, rise-fall). It also confirms teaching observations that tone complexity impedes its correct labelling. However, even in the category of simple tones, the performance was not uniform. While ‘fall’ had high identification accuracy of 78%, the ‘rise’ tone had accuracy of only 54%. It is surprising considering the fact that structurally those tones are the same, differing only in the direction of change. What is more, ‘rise’ was misidentified as ‘fall’ as frequently as 38% of the time, however the opposite pattern of misidentification – ‘fall’ misidentified as ‘rise’ – was found in only 18% of the trials. As predicted, complex tones had the lowest identification rate with considerable inter-tone misidentification. ‘Fall-rise’ was mislabelled as ‘rise-fall’ in almost 40% of the trials and ‘rise-fall’ was mislabelled as ‘fall-rise’ 31% of the time. All this shows that the listeners were relatively accurate in identifying tones as simple or complex, however they had considerable difficulties with identifying a given trajectory of the tone pattern.

The phonetic training in prosody was found to contribute significantly to tone identification. The listeners after the training identified correctly 64% of the tokens compared to the listeners prior to the training who correctly identified 48% of the tones. On the other hand, learning a tone language as another foreign language did not improve tone identification. Both the learners of Chinese and the learners of other non-tone languages performed similarly. At first glance, it may indicate that the phonetic training in prosody increases the accuracy of tone labelling and learning a tone language does not. However, the analysis of interaction between phonetic training and a second foreign language revealed an interesting coupling between phonetic training and learning a tone language. While both groups of listeners (Chinese/non-Chinese) had a similar accuracy rate prior to the phonetic training (in fact the Chinese group was even less accurate than the non-Chinese group), it was the Chinese group that benefited significantly from phonetic training and not the non-Chinese group. The interaction revealed that the significant effect of phonetic training on tone identification was contributed to by the Chinese group. In the case of the non-Chinese group, the phonetic training slightly improved the identification, however the effect was not significant. All this leads to the conclusion that the phonetic training in prosody together with learning a tone language in combination significantly enhance the accuracy of tone identification and labelling. It is not surprising considering the fact that tones are introduced from the very beginning of teaching Chinese, since they are lexically contrastive. Accordingly, it may be assumed that the Chinese group had had much more experience and training in tone identification. In the light of the current results it may be said that the Chinese group outperformed the non-Chinese group due to additional experience with tone identification. Although learning Chinese and pure phonetic training in tone identification are different in nature, the current results show that in combination they increase the ability to identify and label tone patterns. The results also showed that mere phonetic training is not sufficient to significantly improve tone identification, as demonstrated by the performance of non-Chinese groups prior and post phonetic training.

Unlike identification, the discrimination of the same tone patterns was almost perfect with the accuracy level of 92%. There were no significant differences resulting from the complexity of the tones. Both simple and complex tones had similarly high discrimination rate. Neither the phonetic training nor Chinese as a second foreign language were significant predictors in discrimination. This is evidence that psychoacoustic processing of tone patterns is almost perfect despite the lack of experience with tone labelling. It is also evidence that the observed difficulties with tone labelling do not have their source in the lack of sensitivity to either the structure or the direction of the tone trajectory. In other words, the listeners perceived highly accurately that the four tested tone patterns were different and they could accurately match them in a triad of consecutive tones. Discrimination is claimed to tap purely acoustic rather than phonetic or phonological processing (Werker and Logan 1985), while identification relies on categorical perception. According to Logan et al. (1991), discrimination exploits rapidly fading sensory information processed in short-term memory. Correct identification, on the other hand, requires phonetic codes stored in long-term memory. Comparing the performance in identification and discrimination in the current study, we may safely conclude that any difficulties with tone identification do not result from the listeners' inability to perceive tone patterns as different but rather emerge later at the level of conscious matching of the auditory signal with the phonetic label. The inter-stimulus-interval of 500 ms used in the discrimination task strengthened the effect of purely psychoacoustic perception by our listeners. Longer intervals of e.g., 1500 ms activate phonological processing and reduce the ability to discriminate stimuli for which listeners do not store phonological categories (Werker and Tees 1984). We may speculate that lengthening the inter-stimulus-interval would have reduced the discrimination accuracy of the tested tones. Burnham and Francis (1997) reported that increasing the inter-stimulus-interval from 500 to 1500 ms reduced the discrimination of Thai lexical tones by native speakers of Australian English.

Although we have no knowledge of the similar experimental studies investigating the perception of tone patterns used in phonetic training in prosody of English, we may attempt to relate the current results to the observations from teaching the perception of tone contrasts in tone languages. It is frequently reported by teachers of Mandarin Chinese that tone contrasts are one of the most difficult aspects to learn for speakers of non-tone languages (Bluhme and Burr 1971; Kiriloff 1969; Lin 1985). However, even two weeks of tone training in learning Chinese significantly improves identification of those tones (Wang et al. 1999). Although learning pitch patterns to signal lexical contrasts in tone languages and learning the phonetic labelling of pitch patterns seem to be two different tasks, they both share the necessity to form and store tone categories in long-term memory. Our results suggest that the combination of these two tasks may be the most effective method of training the identification of tones, as demonstrated by the performance of the listeners after phonetic training and with Chinese as another foreign language. As shown in our results, the discrimination sensitivity for tone patterns is very high even if listeners have no experience with tone labelling. It leads to the conclusion that the reported difficulties with teaching tone patterns in phonetics are mostly a result of the inability to form and store a phonetic label for each tone.

References

Bluhme, H., and R. Burr 1971. An audio-visual display of pitch for teaching Chinese tones. *Studies in Linguistics* 22: 51-57.

Boersma, P. 2001. Praat, a system for doing phonetics by computer. *Glot International* 10: 341-345.

Bradford, B. 1988. *Intonation in context*. Cambridge: Cambridge University Press.

Brazil, D. 1994. *Pronunciation for advanced learners of English*. Cambridge: Cambridge University Press.

Brazil, D. 1997. *The communicative value of intonation in English*. Cambridge: Cambridge University Press.

Brazil, D., M. Coulthard and C. Johns 1980. *Discourse intonation and language teaching*. Harlow: Longman.

Burnham, D., and E. Francis 1997. The role of linguistic experience in the perception of Thai tones. In A. Abramson (ed.), *Southeast Asian Linguistic Studies in Honour of Vichin Panupong*, 29-47. Bangkok: Chulalongkorn University Press.

Cauldwell, R., and M. Allen 1997. *Phonology*. Birmingham: The University of Birmingham.

Cook, V.J. 1968. *Active intonation*. London: Longman.

Dalton, C., and B. Seidlhofer 1994. *Pronunciation*. Oxford: Oxford University Press.

Grabe, E., G. Kochanski and J. Coleman 2005. The intonation of native accent varieties in the British Isles: Potential for miscommunication? In K. Dziubalska-Kołaczyk and J. Przedlacka (eds), *English pronunciation models: A changing scene*, 311-338. Bern: Peter Lang.

Grabe, E., B.S. Rosner, J.E. Garcia-Albea and X. Zhou 2003. Perception of English intonation by English, Spanish, and Chinese listeners. *Language and Speech* 46: 375-401.

Green, D.M., and J.A. Sweets 1966. *Signal detection theory and psychophysics*. New York: Wiley.

O'Connor, J., and G. Arnold 1973. *Intonation of colloquial English*. London: Longman.

Kelly, G. 2000. *How to teach pronunciation*. Harlow: Pearson Education Limited.

Kiriloff, C. 1969. On the auditory perception of tones in Mandarin. *Phonetica* 20: 63-67.

Lin, W.C.J. 1985. Teaching Mandarin tones to adult English speakers: Analysis of difficulties with suggested remedies. *RELC Journal* 16: 31-47.

Logan, J.S., S.E. Lively and D.B. Pisoni 1991. Training Japanese listeners to identify /r/ and /l/: A first report. *Journal of the Acoustical Society of America* 89: 874-886.

Macmillan, N.A., and C.D. Creelman 1991. *Detection theory: A user's guide*. Cambridge: Cambridge University Press.

Ranalli, J.M. 2002. *Discourse intonation: To teach or not to teach?* Birmingham: The University of Birmingham.

Roach, P. 1991. *English phonetics and phonology. A practical course*. Cambridge: Cambridge University Press.

Setter, J. 2005. Communicative patterns of intonation in L2 English teaching and learning: The impact of discourse approaches. In K. Dziubalska-Kołaczyk and J. Przedlacka (eds), *English pronunciation models: A changing scene*, 367-389. Bern: Peter Lang.

Stanislaw, H., and N. Todorov 1999. Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers* 31(1): 137-149.

Taylor, D.S. 1993. Intonation and accent in English: what teachers need to know. *International Review of Applied Linguistics in Language Teaching* 31(1): l-22.

Wang, Y., M.M. Spence, A. Jongman, J.A. Sereno 1999. Training American listeners to perceive Mandarin tone. *Journal of the Acoustical Society of America* 106: 3649-3658.

Werker, J.F., and J. Logan 1985. Cross-language evidence for three factors in speech perception. *Perception and Psychophysics* 37: 35-44.

Werker, J.H., and R.C. Tees 1984. Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America* 75: 1866-1878.

Underhill, A. 1994. *Sound foundations*. Oxford: Heinemann.