

The effect of F0 peak alignment, height and segmental base on sentence type perception across languages

Veronika Makarova

Meikai university, AIST, Japan

E-mail: makarova.veronika@aist.go.jp

ABSTRACT

This paper addresses the effect of F0 peak alignment, F0 peak height and segmental base on the perception of sentence type of 3-syllable one-word utterances by Japanese and Russian listeners. The reported study was performed with re-synthesized stimuli which differed by 14 locations (alignments) and 2 heights of F0 peak and two segmental bases. The stimuli were produced by modifications of re-synthesized original natural speech declarative (declarative base) and interrogative (interrogative base) utterances. Sentence type identification of randomized stimuli as declarative, interrogative and exclamatory by Japanese and Russian subjects showed that pitch peak height has a strong influence on the perception of both groups of subjects, base has almost no effect, whereas F0 peak alignment affects the perception of Russian subjects significantly more than of Japanese subjects. Alignment shifts towards the end of the accented vowel cause an increase in interrogative judgments by Russian subjects.

1. INTRODUCTION

1.1. Intonation functions

Intonation performs a number of important functions in human speech. I suggest distinguishing between the information-structure function (i.e. it identifies the structure of spoken text, such as beginnings and ends of information units and their hierarchies, signifies the information focus, new and given information), grammar-syntactic (it surfaces some parts of syntactic and grammatical components of the spoken message), pragmatic-discoursal (suggests turn-taking and renders communicative intentions), semantic (expresses emotive and attitudinal meanings), speaker identification (reflects some individual speaker-dependent features, such as speech rate, F0 range and voice source characteristics, etc), esthetical (in speech-related arts), psychological (such as the establishment of mother-child bonds, or attempts to captivate the minds of crowds in politicians' or clergymen speeches), and psychotherapeutic (speech-related psychotherapy, prayers, chants, etc.). It has been earlier proposed that intonation functions can be divided into two subsets, the structure-component mode (alias, 'linguistic' functions or 'intonation grammar'), to which most authors

would assign the first three items from the above given list, and the speaker-pragmatic mode (alias, 'paralinguistic' function or 'phonetic implementations'), which is usually associated with the next two functions (semantic, speaker-identification) from the list [1,2,3]. The last four functions from the above list are ignored by most researchers, but can be tentatively placed into the speaker-pragmatic mode.

1.2. Intonation features involved in sentence type distinctions

One of the best known manifestations of the grammar-syntactic function of intonation is the ability of intonation to render sentence type distinctions. This function is so common across languages that it is often considered one of linguistic universals [4]. In regard of the exact intonation-syntactic type links, it was observed in particular that despite the existence of numerous exceptions [5], 'yes/no' interrogatives in many languages tend to be associated with the rising or high pitch, whereas declaratives tend to have a falling tonic pitch movement [2]. Correspondingly, in speech perception, a rising (or rising-falling) tonic elicits more 'question' judgments, whereas a falling tonic generates more 'declarative' judgements [6].

1.3. The alignment of pitch events

Pitch event alignment can be defined as the exact location along the segmental continuum where some important changes in pitch movements (peaks, troughs, onsets and offsets) occur. Some recent studies of a few unrelated languages suggest that besides direction and height of pitch movements, their alignments may also contribute to the expression of sentence type contrasts [7]. It is not yet clear what effect alignment shifts may have on the perception of speakers of other languages and whether and in what ways alignment can be connected with other prosodic parameters. To investigate this issue, I undertook an experimental study described below.

1.4. Intonation markers of sentence types in Japanese and Russian

In general, Japanese has lexical markers for sentence types, which help to distinguish 'yes/no' interrogatives and exclamations from declaratives. However, in frequently found elliptical sentences, intonation serves as the major tool for sentence type identification. Declaratives typically

have a terminal fall, whereas yes-no questions – a rise. Many patterns can be found in exclamations. Alignment of pitch events was reported to be used, in particular, for differentiating some types of questions [8].

In Russian, 'yes-no' questions, declaratives and exclamations typically do not bear any lexical or grammatical markers of sentence types, and can only be disambiguated in perception via intonation cues. In some cases (such as one-word three syllable sentences) the three sentence types can have very similar rising-falling patterns, yet they are differentiated by the speakers of Russian with very high accuracy. Some recent production experiments have shown that these sentence types have slight differences in the alignment of the pitch peak regarding the accented vowel onset-offset [6,9]. It was possible to expect that these differences are utilized in perception as well.

Earlier comparisons of the perceptual strategies of Russian and Japanese listeners reveal both some similarities related to the perception of pitch height and slope, but difference in labeling of sentence type [6].

2. MATERIALS AND METHODS

The experiment was conducted with a set of re-synthesized utterances with 14 different locations of pitch peak. One-word three-syllable utterance 'banany' [*bananas*] with the lexical stress on the penultimate syllable was chosen as the basis for re-synthesis in the experiment. The word was read as a declarative by one female speaker of standard Russian aged 29 at the time of the recording. The recording was sampled at 16 KHz.

Re-synthesis of the original stimuli and their subsequent manipulations were performed with Praat software programme (PSOLA analysis and resynthesis mode). To test the effect of height on listeners' perception of pitch peak alignment contrasts, I re-synthesized the utterance with two different values of F0 peak: 320 Hz (referred to as 'high peak') and 270 Hz ('low peak'). To investigate some possible effect of segmental base (i.e. some segmental quality characteristics) on the perception of sentence type, all stimuli were re-synthesized with identical modifications performed on 2 segmental bases: the one of the original declarative (declarative base) and interrogative (interrogative base) utterances. Four sets of stimuli were prepared with 14 identical modifications of pitch peak alignment by modifying the original F0 values of the source utterances. The peak was shifted within the accented and postaccented syllables and in about 20ms intervals. The above described manipulations yielded the total of 56 stimuli sound files (2 heights X 2 bases X 14 alignments). All the resulting sound files were copied 5 times each with 2.5s silent intervals between them and randomized to obtain the sound record for the experimental session (the total number of re-synthesized modified speech signals used for the perceptual experiment was therefore 280).

3. RESULTS OF SENTENCE TYPE IDENTIFICATION BY SUBJECTS

Factor ANOVA designs were conducted to test the perception data for the effects of alignment, subject group, and pitch height. The results show that alignment is strongly significant (at $p < 0.001$) for stimuli identification by Russian subjects, but is insignificant for the Japanese subjects. The factor of pitch height is significant for the perception of both groups of subjects. The factor of base is only significant for the perception of utterances as exclamatory and interrogative by Russian subjects (at $p < 0.01$). There is a significant difference (at $p < 0.001$) between the perception of stimuli across the subject groups.

The results of the experimental stimuli identification by subjects are represented in graphic form in Figures 1-8.

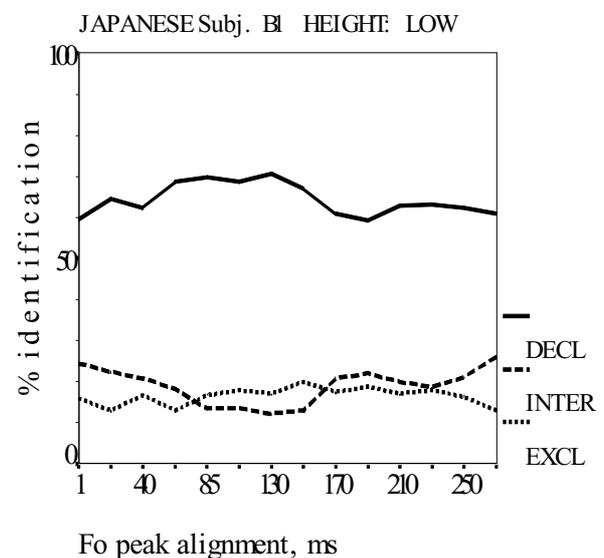


Figure 1. Perception of lower pitched stimuli with shifted peak alignment by Japanese subjects (declarative base)

As indicated by the graphs, pitch height affects the subjects' perception: lower pitch peak height yields predominantly declarative judgments by both groups of speakers, whereas higher pitch yields an increase in the percentage of non-declarative judgments. Second, while alignment does not in the least affect the perception of Japanese subjects in the lower peak set, the rightward alignment shift very slightly increases the number of exclamatory judgments in the higher pitch set. For Russian subjects, the rightward shift of pitch peak alignment causes a decrease of declarative (in the lower set) and exclamatory (in the lower set) judgments, and an increase of interrogative judgment. Third, for Russian subjects, the effect of alignment is stronger for the set of higher-pitched stimuli where we observe a clear categorical shift (located close to the accented vowel offset) in perception from predominantly interrogative to predominantly exclamatory.

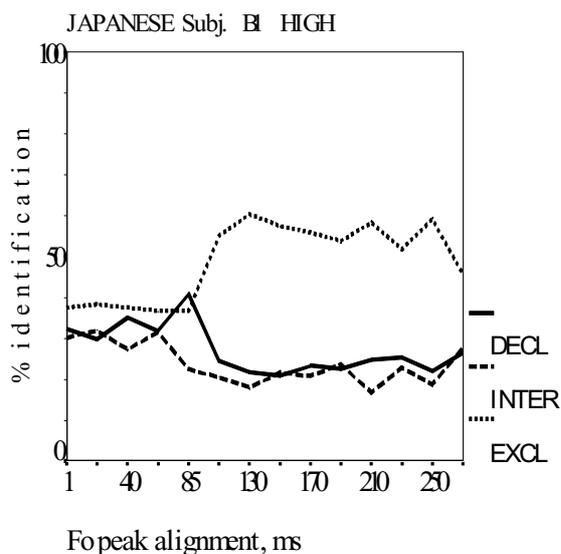


Figure 2. Perception of higher pitched stimuli with shifted peak alignment by Japanese subjects (declarative base)

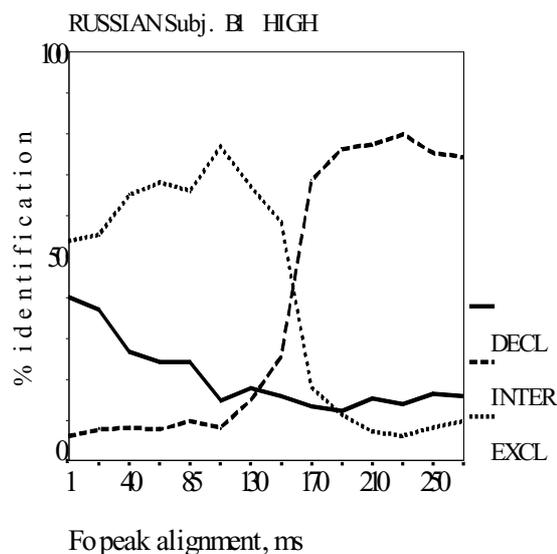


Figure 4. Perception of higher pitched stimuli by Russian subjects (declarative base)

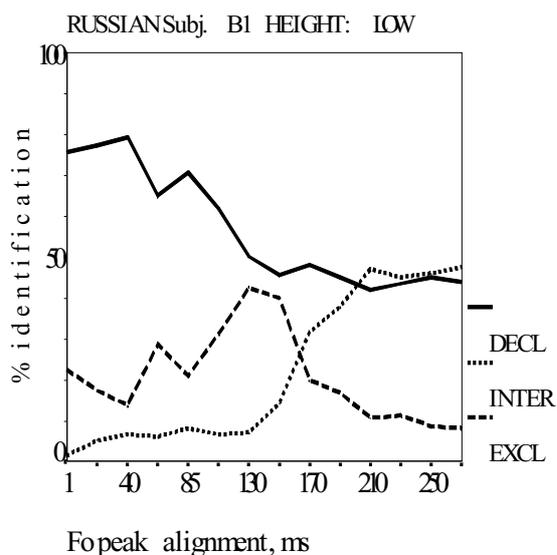


Figure 3. Perception of lower pitched stimuli with alignment shift by Russian listeners (declarative base).

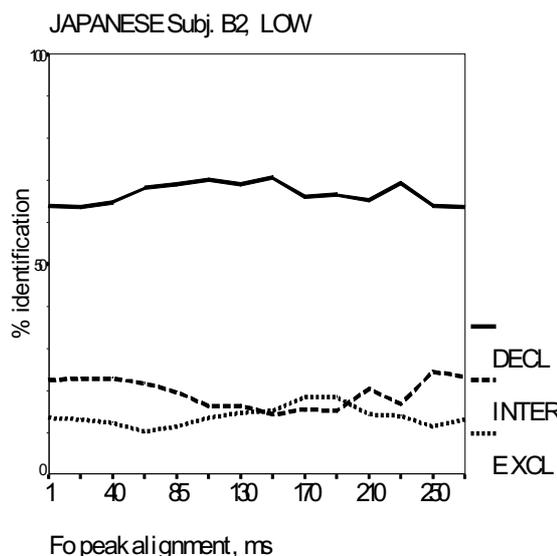


Figure 5. Perception of lower pitched stimuli by Japanese subjects (interrogative base)

The 'best declarative' judgements in all sets lie at the earliest alignments, which agrees well with the production data according to which declaratives have early alignments. The highest percent of exclamatory identification is found for the Russian subjects, close to the accented vowel onset. However, Japanese subjects, give more 'exclamatory' labels to stimuli with delayed pitch peak. Interrogative judgments hardly appear at all in the responses by Japanese subjects.

4. CONCLUSION

We see from the results of the reported experiment that a prosodic feature (pitch peak alignment) systematically employed in speech production to produce certain semantic values (grammatical meanings of sentence types) affects the listeners' perception (in case of Russian), whereas it has no or very little impact on the perception of subjects whose language (Japanese) does not use this feature in production.

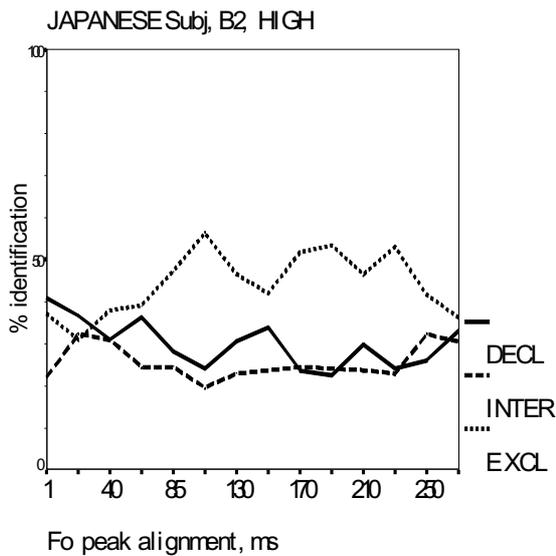


Figure 6. Perception of higher pitched stimuli with alignment shift by Japanese subjects (interrogative base)

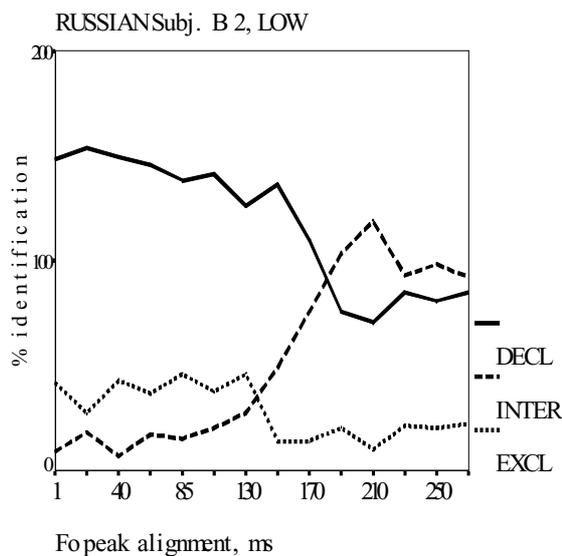


Figure 7. Perception of lower pitched stimuli with alignment shift by Russian subjects (interrogative base).

The experiment confirms the existence of a threshold of declarative/non-declarative perception for Russian subjects [7]. The change in perception occurs earlier and more pronounced in higher pitch sets. It should be noted that the location of the interrogative perception boundary (20-40 ms from the accented vowel offset) is close to the average values of the parameter in speech production [7]. A bigger ratio of interrogative and exclamatory judgments in the set with higher peak confirms the results of previous studies which connect higher pitch with interrogativity and/or emphasis [3, 14]. This experiment displays the possibility of links between pitch height and alignment for the speakers of some languages.

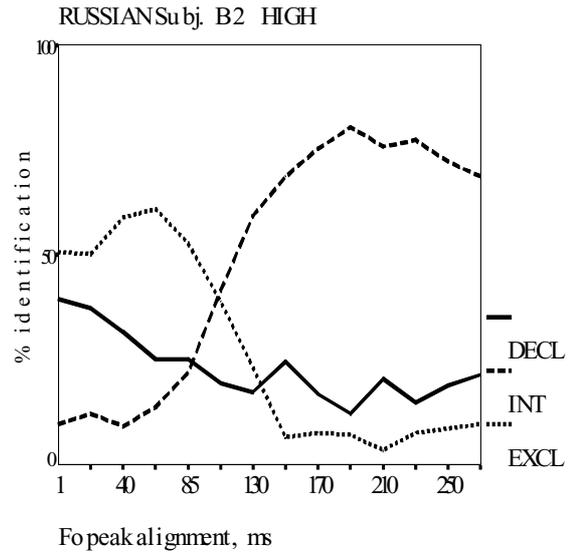


Figure 8. Perception of higher pitched stimuli with alignment shift by Russian subjects (interrogative base).

REFERENCES

- [1] Campbell, N. 2002. Towards a grammar of spoken language: Incorporating paralinguistic information. *ICSLP 2002*, 673-676.
- [2] Ladd, D. R. 1996. *Intonational phonology*. Cambridge: Cambridge University Press.
- [3] Gussenhoven, C. 2002. Intonation and interpretation: Phonetics and phonology. *Speech Prosody*, 11-13 April 2002, Aix-en Provence, France.
- [4] Bolinger, D. 1978. Intonation across languages. In J. P. Greenberg, C. A. Ferguson & E.A. Moravcsik (eds) *Universals of human language*, Phonology 2, 471-524, Stanford, Stanford University Press.
- [5] Gussenhoven, C. & Chen, A. 2000. Universal and language-specific effects in the perception of question intonation. *ICSLP 6*, 91-94.
- [6] Makarova, V. 2001. Perceptual correlates of sentence type intonation in Russian and Japanese. *Journal of Phonetics*, 29 (2), 137-154
- [7] Gosy, M. & Terken, J. 1994. Question marking in Hungarian: timing and height of pitch peaks. *Journal of Phonetics*, 22, 269-281.
- [8] Abe, I. 1998. Intonation in Japanese. In: Hirst, D. & Di Cristo, A. *Intonation systems: A survey of twenty languages*. Cambridge: CUP, pp. 360-375.
- [9] Makarova, V. 2001. Phonetic approaches to issues in Russian intonation. In: Zybatow, G., Juhghanns, Mehlhorn, G., Szucsich, L. (Eds.). *Current Issues in Formal Slavistics*. Peter Lang, 49-58.
- [10] Ladd, D.R. & Morton, R. 1997. The perception of intonation emphasis: continuous or categorical. *Journal of Phonetics*, 25, 313-342.