Non- and Quasi-lexical Realizations of "Positive Response" in Korean, Polish and Thai

Maciej Karpiński, Janusz Kleśta, Emilia Szalkowska

Institute of Linguistics
Adam Mickiewicz University, Poznań, Poland
{Maciej.Karpinski|Janusz.Klesta|Emilia.Szalkowska}@amu.edu.pl

Abstract

In the present paper, various categories of "positive response" in Korean, Polish and Thai task-oriented dialogues are examined. Pitch contours are categorized using visualizations and close listening. Basic F_0 parameters, including F_0 direction change, range and rate are measured and analyzed for selected sets of expressions. The results are discussed in pragmaphonetic terms. Language-specific and universal aspects of intonation are stressed.

1. Introduction

This paper presents a basic comparative study of Korean, Polish and Thai short words, quasi-words and other vocalizations used to perform the dialogue moves ([1] [2]) collectively referred to as "positive responses" [3]. The positive responses in question may refer to both the content of a preceding utterance as well as its occurrence alone. These two pragmatic meanings are often difficult to separate.

Some of the abovementioned expressions are produced as non-linguistic vocalizations with a very hazy segmental structure, while others are "fully legitimate" lexical entities [3]. Although hard to contain within one category, they show certain common properties. They may occur in a number of discourse functions, depending on the context and the form of realization ([4] [5]). They are frequently produced as backchannels ([6] [7] [8] [9]), but they may also form "standard" dialogue turns or act as discourse markers, organizing the turns internally or steering the flow of conversation and turn-taking ([10] [11] [12]). Their prosodic shape may not only convey information about the dialogue move category, but also signal emotional and attitudinal meanings [13]. Therefore, these expressions play an important part not only on the phatic or semantic but also on the emotional level of dialogue.

The three languages under study come from typologically different families. Korean is agglutinative, Polish – inflectional, and Thai – isolating (analytical). Moreover, Thai is tonal, while Polish and Korean – intonational (the latter being historically tonal). The cultural and geographical distances between Korea, Poland and Thailand also result in different communicational behaviors of the respective countries' inhabitants.

2. Data acquisition and inventory overview

The data under study come from *Pol'n'Asia Map Task Corpus* and *PoInt Corpus* (comp. [14], [15], [16]), based on the same non-facing recording procedure for the speakers of the three languages. Speech was recorded digitally on two channels, but the speakers were not acoustically isolated from each other.

Only non-overlapping units of speech were extracted for further analyses. Consequently, typical overlapping backchannels have been excluded. Distorted, low quality, laryngalized or whispered utterances were also omitted in the automatic analyses. For each language, the targets were grouped into categories on the basis of their segmental structure and dialogue function.

366 Korean target 'positive responses' were extracted from 10 map task dialogues lasting from 3'6'' to 8'19'' (50'42'' in total). The material was realized by 6 female pairs, 1 male pair and 3 mixed pairs. The number of targets per dialogue ranged between 12 and 77 (excluding responses overlapping with the interlocutor's turn). Although the extracted units were assigned to 11 categories, only the five most numerous groups were analyzed as they accounted for 91,5% of the whole inventory.

It was found that the choice of units used by each pair strongly depended on their degree of intimacy, social position, and age, remaining different for the formal speech level (*kyŏksikch'e*, where *ne* was used), and a low style (*panmal*) of the informal speech level (*pikyŏksikch'e*, where *ne* was virtually absent). Other common native Korean expressions like *kŭraeyo* (*so it is*), *maja* ([*that's*] *right*) or *chohayo* ([*that's*] *good*, *all right*) did not occur at all (exception *kŭrae* – 1 occurrence), which may be a task-specific phenomenon. Among non-native expressions, the /mm̥m/–/yhm/ group outnumbered the others. Borrowed expressions in total accounted for 12% of the inventory. The usage of particular expressions was also strongly speaker-specific: For example, 73% of all /mm̥m/–/yhm/ items in the corpus were produced by one speaker only, 100% of /ə/ by another participant, while 67% of *ye* by yet another person. Only *ŭng* had a quite regular distribution across almost all recordings with the exception of high speech level dialogues.

Table 1. *Expressions of 'positive response' in Korean*

No	category	occurrences	mean F_0 change	maximum F_0 change
1	/ɯŋ/, /ɯm/	123	16%	38%
2	/ɔ/, /ʌ/	60	15%	31%
3	/a/	52	15%	41%
4	other vowels, mostly /ə/	9	-	-
5	/ne/	68*	17%	42%
6	/je/	9	-	-
7	/mm̥m//yhm/	33	17%	33%
8	/aha/	6	-	-
9	/ok^{h}ej/	4	-	-
10	/jes/	1	-	-
11	/kɯɾe/	1	-	-

*) Only for /ne/, the proportion of occurrences as "yes-response" was substantial enough for separate analyses.

For Polish, the analyzed units were extracted from a two-hour corpus of ten map task dialogues with equal numbers of female and male speakers. The session durations ranged from 6'30'' to 19'49''. Ten categories of units were distinguished among various short realizations of "positive responses". The five most numerous categories, accounting for 85% of the tokens, were selected for further analyses.

Table 2. *Expressions of 'positive response' in Polish*

No	category	occurrences	mean F_0 change	maximum F_0 change
1	/axa/	80	22%	63%
2	/ax/	8	-	-
3	/dobra/	86	20%	57%
4	/dobʒe/	50	-	-
5	/mm̥m/	139	22%	61%
6	/no/	114	15%	57%
7	/nodobʒe/	12	-	-
8	/nodobra/	17	-	-
9	/okej/, /okiː/	32	-	-
10	/tak/	238*	19%	68%

*) Only for /tak/ was the proportion of occurrences as "yes-response" substantial enough for separate analyses.

While all of the listed units may be used for positive response and function as „acknowledgement", some of them have more specific applications. /axa/ seems to be predestined to express acknowledgement with a certain amount of surprise, while /ax/ is a bit old-fashioned and quite rare; /dobra/ and /dobʒe/ have similar usages and neutral meanings, with the former being less formal (both can form quasi-phrases with /no/); /mm̥m/ is a nasal vocalization sounding similarly to English *mhm*, but bearing somewhat different meanings; /no/ may occur in a number of meanings, including "plain" acknowledgement, the expression of surprise, admiration but also disappointment; the origin of /okej/ is obviously English and its meaning is not far from the original, while /tak/ seems to be the closest equivalent of English *yes*, having, nevertheless, other meanings or usages, too (e.g., in exophora). While all the expressions sound relatively polite, only two of them (/tak/, /dobʒe/) seem to be acceptable in a formal setting.

The analyses carried out for the Thai language were based on a recording corpus including a set of 10 map task dialogues realized by three male and seven female pairs. The duration of the individual recording sessions ranged from 2'28'' to 7'18'', averaging 4'28" per session. The corpus of Thai speech enabled the extraction of 339 non-lexical words, quasi-words and nasal vocalizations, constituting "positive responses". Although the extracted items included expressions representing 11 different categories, only the four most numerous groups, accounting for nearly 92% of the whole material, were analyzed within the current study.

Approximate meanings of the targets are as follows: /t͡ɕʰaj/ - non-lexical word equivalent to English *yes*, *that is so*, *that is correct*; /t͡ɕʰajlɛːw/ - emphatic non-lexical word equivalent to English *that's right*, *exactly, exactly so* (an extended form of the former); /əː/ - quasiword equivalent to English *yeah;* /mː/ nasal vocalization equivalent to English *mhm*; /oːkʰeː/ borrowed western non-lexical word equivalent to English *okay;* /aha/ borrowed western quasiword equivalent to English *uh-huh;* /mm̥m/ borrowed western nasal vocalization equivalent to English *mhm;* /kʰrapˀ/ - particle of acknowledgement equivalent to English *yes*; /kʰrapʰom/ - emphatic particle of acknowledgement equivalent to English *yes sir*; /ɔː/ - quasiword expressing statement's acknowledgement with slight surprise equivalent to English *oh*, *aha*; /mː/ - nasal vocalization expressing statement's acknowledgement with slight uncertainty equivalent to English *hmm* (characterized by a very steep pitch fall).

Table 3. *Expressions of 'positive response' in Thai*

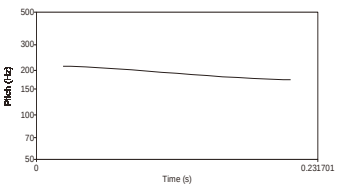
No	category	occurrences	mean F_0 change	maximum F_0 change
1	/t͡ɕʰaj/	46	16%	40%
2	/t͡ɕʰajlɛːw/	1	-	-
3	/əː/	90	9%	37%
4	/mː/	113	13%	41%
5	/oːkʰeː/	9	-	-
6	/aha/	6	-	-
7	/mm̥m/	61	11%	42%
8	/kʰrapˀ/	2	-	-
9	/kʰrapʰom/	2	-	-
10	/ɔː/	7	-	-
11	/mː/	2	-	-

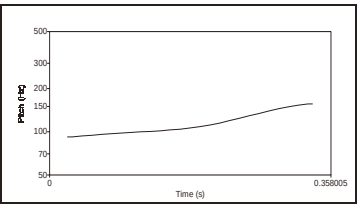
3. Acoustic analysis of selected expressions

Basic pitch parameters were extracted from the signals automatically using Praat [17]. Simple measures of pitch change direction, range and rate were defined. Auditory and visual analysis were used to find typical pitch contours and to search for certain less obvious tendencies in the data set.

The typical affirmative intonational contours of native Korean informal /ɯŋ/, /ɯm/, /ɔ/, /ʌ/ and /a/ expressions show an overall falling tendency, while for the non-native /mm̥m/ expression, a rising contour was applied in 97% of the realizations (see tables below). Instead of just one prevailing melody, two contours (falling and rising) with similar distribution (respectively 38% and 43%) were observed for /ne/. Although the falling contour transpired to be the dominant realization of /ne/ as the yes-answer to a polar question (58%), the rising one appeared to be more common in the acknowledgements (53%). More complex rising-falling contours of /ne/ were not analyzed here because of their less frequent occurrence.

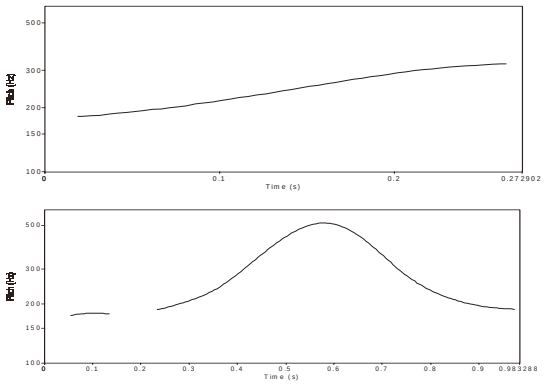
The average percentage pitch change for the groups of the most typical contours in each of the analyzed expression categories ranged between 15% and 17% (see Table 1). The ranges of values calculated for individual categories overlapped considerably, with a value of 42% being the upper limit. It seems that this limiting value is exceeded by neither speakers of Korean nor users of Thai (see further on).

expression category (% of expressions in category)	**typical intonational contour**
/ɯŋ/, /ɯm/ **(91%)**	
/ɔ/, /ʌ/ **(68%)**	
/a/ **(79%)**	
/ne/ **(38%)**	

/mm̥m/, /yhm/ **97%**	
/ne/ **43%**	

The shapes of pitch contours in Polish varied extensively for all of the analyzed units. Unlike for Korean or Thai, it was impossible to select one typical intonational realization for a given expression here. To obtain a more detailed view of possible intonational contours, the realizations of /axa/ were analyzed more closely. /axa/ was chosen because nearly all of its occurrences expressed "acknowledgement", forming the most homogeneous group. Seven pitch contour categories were observed for this expression: a plain rise (14); a plain fall (28); a fall preceded by a small rise (11); a rise preceded by a small fall (13); a flat contour (2); a fall-rise (9) and a rise-fall (2).

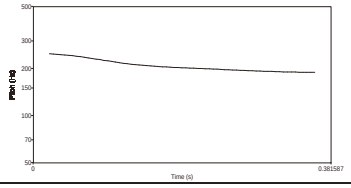
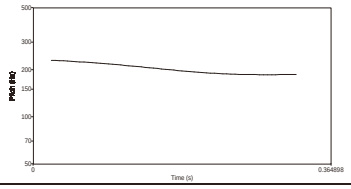
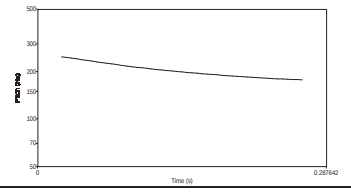
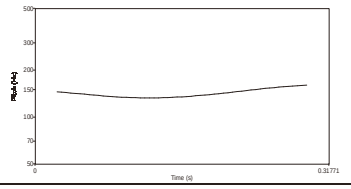
Since the shapes of intonational contours varied substantially for each type of expression, forming numerous subcategories, it was difficult to put forward any testable hypotheses about the shape-function relationship. Considering the most general and not very precise criterion of pitch direction that takes into account only the order of F_0 minimum and maximum on the time axis, only for /mm̥m/ rising contours proved to be clearly predominant (76%). The proportion of rising contours for /axa/ was 55%, while it reached 35% for /dobra/, 49% for /no/, and 63% for /tak/. However, the application of more detailed distinctions between dialogue moves (e.g., instruction acknowledgement vs. statement acknowledgement) showed that for /axa/, /no/, /tak/ and /dobra/, the mean value of relative pitch change was meaningfully higher when they functioned as instruction acknowledgements than in the case of statement acknowledgements. This coincides with the fact that rising contours were generally more frequent in instruction acknowledgements. Certainly, some factors going far beyond dialogue move categories must be responsible for the observed variation, at least along certain dimensions. Below, an exemplary realization of /axa/ with a standard rising contour is contrasted with an emotional rising-falling trace.



The distribution of the pitch change coefficient (a simple measure of direction and steepness for non-complex F_0 contours) was close to normal only for /mm̥m/, while for /dobra/, /tak/ and /axa/ the distributions were closer to binomial, which resulted from a trace number of flat contours.

The mean value of pitch change range for Polish expressions varied from 15% to 23% (see Table 2), with the highest result for /axa/. In most cases, it was higher for rises than for falls. The difference between the values for /tak/ as a "yes-answer" and a realization of other types of "positive response" is worth noticing. As was the case with /tak/ as a "yes-answer", /no/ also featured a very limited variation range. The mean value of pitch change for all the items examined here reached 22%.

Typical, i.e. the most frequent intonational contours for each of the Thai expression categories under study are presented in the table below, while a detailed discussion of the remaining contours realized with the expressions under investigation will be the subject of a separate paper because it exceeds the limited scope of this comparative study.

expression category (% of expressions in category)	typical intonational contour
/t͡ɕʰaj/ **(91%)**	
/əː/ **(79%)**	
/mː/ **(89%)**	
/mm̥m/ **(74%)**	

As can be seen in the above figures, the typical intonational contours had a predominantly falling orientation in all of the three original Thai expression categories under study, while a distinctly rising contour was commonly applied in the case of the borrowed western affirmative nasal vocalization /mm̥m/. We might, therefore, hypothesize that the quasi-word /əː/ as well as the nasal vocalization /mː/, which are both originally Thai affirmative expressions, are merely reduced, simplified versions of the basic Thai affirmative non-lexical word /t͡ɕʰaj/, and are consequently realized with a similar F_0 contour, while the borrowed expression of /mm̥m/ takes on a completely different intonational realization, being a reduced version of an affirmative word in a different language. We cannot forget, though, that in the case of Thai the intonational contour produced with a mono-syllabic intonational phrase, resembling those under analysis, is always the combinatory effect of the phrasal intonation pattern being superimposed on the lexical intonation, i.e. the lexical tone. Therefore, as separating the

two intonational effects is an extremely difficult task, it remains an open question whether the presumably reduced versions of the basic affirmative word involve copying its lexical tone or the application of the phrasal intonation pattern it is usually produced with, especially that the F_0 curves characteristic of the lexical tones connected with the articulation of the analyzed expressions according to their orthographic transcription slightly differ from the F_0 traces we could observe in their realizations.

The average percentage pitch change for the expressions under study ranged from 9% for /əː/ to 16% in the case of /t͡ɕʰaj/. The average values of the above parameter along with the maximum values observed for expressions from the individual categories are presented in Table 3.

4. Universal and language-specific features of the analyzed units

As can be seen in the presented figures, the average frequency of occurrence for the analyzed expressions of 'positive response' in the recorded material was not only quite substantial but also largely similar for all of the three languages under investigation, ranging between 6.5 items per minute for Polish and 7.5 items per minute for Thai, with the intermediate value of 7.2 items/minute for Korean. Also, the overall number of expression categories constituting the inventories of individual languages was almost identical (10 for Polish, 11 for Korean and Thai), with, consistently, four (Thai) or five (Korean and Polish) most popular of them accounting for about 90% of all the analyzed tokens.

However, while positive responses in Korean and Thai seem to show relatively distinct tendencies in the area of intonational contours realized with individual expression categories, their Polish equivalents transpire as much less consistent in this respect, featuring a wide range of different melodies imposed on various productions of the same utterance.

Besides, while the inventories of expressions for the three analyzed languages overlap to some extent due to such obvious borrowings as *okay* (articulated slightly differently depending on the language), it seems to be particularly striking that the nasal vocalization *mhm* (transcribed here as /mm̥m/) not only ranked among the four/five most popular expression categories for each of the languages under study but was also surprisingly consistently produced with a rising contour, regardless of the melody predominantly realized with positive responses in a given language.

Finally, it is also remarkable that while the normalized pitch change, calculated as the difference between the F_0 maximum and minimum divided by the former, was never higher than approximately 40% for positive responses in the two Asian languages, it often reached about 60% in the case of Polish expressions, with accordingly higher average values of this coefficient.

Further research will include a more detailed analysis of the dialogue functions for the units in question as well as extending the scope of study to other languages (Vietnamese).

This study is a part of Pol'n'Asia project, funded by the Council for Science (project code: H01D 006 27).

5. References

[1] Carlson, L. 1983. *Dialogue Games: An approach to discourse analysis*. Boston: D. Reidel.

[2] Carletta, J., Isard, A., Isard, S., Kowtko, J., Doherty-Sneddon, G. and Anderson, A. 1996. *HCRC dialogue structure coding manual*, Human Communications Research Centre, University of Edinburgh, Edinburgh, HCRC TR – 82.

[3] Karpiński, M. 2004. The form and function of selected of selected non-lexical and quasi-lexical units in Polish map task dialogues. [In:] *The Proceedings of SASR Conference*. Kraków.

[4] Schegloff, E. A. 1982. Discourse as an interactional achievement: Some uses of 'uh huh' and other things that come between sentences. [In:] Deborah Tannen [Ed.], *Analyzing Discourse: Text and Talk.* Georgetown University Press, Washington, D.C., pp. 71 – 93.

[5] Ward, N. 2005. Non-Lexical Conversational Sounds in American English, *Pragmatics and Cognition*.

[6] Yngve, V. H. 1970. On getting a word in edgewise. [In:] *Papers from the Sixth Regional Meeting of the Chicago Linguistic Society*, pp. 567 – 577, Chicago: University of Chicago.

[7] Caspers, J. 2000. Melodic characteristics of backchannels in Dutch map task dialogues, [In:] ICSLP-2000, vol. 2, pp. 611 – 614.

[8] Ward, N., Tsukahara, W. 2000. Prosodic Features which Cue Back-channel Responses in English and Japanese. [In:] *Journal of Pragmatics*, vol. 23, pp. 1177 – 1207.

[9] Cathcart, N., Carletta, J., Klein, E. 2003. A model of back-channel acknowledgements in spoken dialogue. EACL Proceedings, pp. 51 – 58.

[10] Schiffrin, D. 1987. *Discourse markers*. Cambridge: CUP.

[11] Swerts, M., Wichmann, A. & Beun, R. 1996. Filled pauses as markers of discourse structure. *The Proceedings of ICSLP*.

[12] Byron, D. K., Heeman, P. A. 1997. Discourse marker use in task-oriented spoken dialog. *The Proceedings of Eurospeech 97*, pp. 2223 – 2226.

[13] Nagano-Madsen, Y., Ayusawa, T. 2005. Prosodic correlates of attitudinally-varied backchannels in Japanese. *Fonetik 2005 Proceedings*. Goeteborg.

[14] Anderson, A. H., Bader, M., Bard E., Boyle, G. E., Doherty, G., Garrod, S., Isard, S., Kowtko, J., McAllister, J., Miller, J., Sotillo, C., Thompson, H., Weinert, R 1991. The HCRC Map Task Corpus. *Language and Speech*, 34(4), pp. 351 – 366.

[15] Aono, M., Ichikawa, A., Koiso, H., Satoh, S. et al. 1994. The Japanese Map Task Corpus.

[16] Grabe, E., Post, B. and Nolan, F. 2001. Modelling intonational Variation in English. The IViE system. [In:] Puppel, S. and Demenko, G. (Eds). *Proceedings of Prosody 2000, 51-57*, Adam Mickiewicz University, Poznan, Poland.

[17] Boersma, P. & Weenink, D. 2005. Praat: doing phonetics by computer (Version 4.3.16) [Computer program]. Retrieved June 22, 2005, from http://www.praat.org/