Is Language Special? Anticipation Timing Accuracy of End of Turns in Known and Unknown Languages

Franziska Schaller[*], Horst M. Müller

Experimental Neurolinguistics Group, Bielefeld University, Germany

Abstract   Structured signaling in the acoustic environment between two individuals usually leads to turns to avoid interference. Turn-taking in human communication is a precise system that enables interlocutors to interact very efficiently. Previous studies have detected criteria that allow for optimized timing within a conversation. For instance, lexico-syntax seems to be of outstanding relevance. Other aspects still under consideration in this context are prosody and rhythm beside others. In the current study, we focused on the question if language carries universal acoustic features which might make turn-taking in human communication uniquely efficient in contrast to e.g. 'turn-taking' in animals. We aimed at getting an impression of how language specific properties other than content and grammatical structure affect anticipation performance. Therefore, we contrasted the Anticipation Timing Accuracy (ATA) for mother-tongue stimuli in German, for items in six foreign languages (English, Italian, Polish, Turkish, Arabic, and Korean) and for simple sinusoidal tones. Results showed significant differences between the ATA of the foreign language stimuli. German subjects anticipated the ends of utterances in Indo-European languages and in stress-timed languages (German, English, Arabic) significantly better than the ends of items in non-Indo-European languages and in syllable-timed languages (Italian, Polish, Turkish, Korean, restrictions apply). We conclude that interlocutors' end-of-utterance anticipation performance is influenced by language inherent universal acoustic features.

Keywords   Anticipation, ATA, Speech Pattern, Spoken Language, Turn-taking

1. Introduction

The efficiency of intraspecific communication is strongly influenced by the respective communication channel. The amount of information mediated in a certain amount of time may thus vary depending on whether a chemical, acoustic, or visual channel is used. Beside the sequentiality of the respective channel, the internal timeline of a dialogue between two individuals is of relevance for the efficiency of communication. Whereas interlocutors may both send visual signals at the same time if a visual channel is used, this opportunity is missing if only an acoustic channel may be used. Complex acoustic signals with duration of up to several seconds require a deferred use of the communication channel in order to avoid interruptions or interferences. Interlocutors therefore attend to a turn structure. In human communication, a turn is not to be confused with a sentence [1]. Rather, a turn is defined as comprising all the utterances of a speaker until the listener takes the conversation over, that is, takes the turn. Turn-taking is thus defined as the transition from one turn to the other, as the transfer of speech from one speaker to the next [2]. This turn-taking may not occur at random points in the speaker's turn but only at *transition relevant places* [1]. In order to achieve high temporal efficiency in communication, listeners have to anticipate these *transition relevant places* and thus the end of a perceived speaker's turn. Human communication thus demands a great deal of interlocutors. However, complex birdsong shows a structured succession of sound signals between two or more individuals as well, which may be categorized as chorus, duet, or antiphony [3].

As suggested by the turn-taking system developed by Sacks et al. [4], dialogue partners show accurate timing when they interact communicatively. This turn-taking system has been repeatedly discussed during the last 40 years. As a result, the projectionists' point of view seems to have been confirmed. From their perspective, the precise speaker changes occurring in everyday conversations are possible because recipients anticipate the end of the speaker's turn and, therefore, know when he or she will stop talking. But what is this anticipation process based on? It seems that lexico-syntactic characteristics of an utterance are of particular importance [5-7]. Additionally, the relevance of some prosodic aspects and suprasegmental

features is supported by the results of a series of studies [8-10]. Wells and Macfarlane [8], Heldner et al. [9], and Dombrowski and Niebuhr [11] found evidence for the relevance of the last major accent and specific F_0-patterns in turn-end anticipation in corpora of different languages (Swedish, English, and German). Other prosodic cues that might be important for a successful anticipation process are the boundary tone of the turn end (Barkhuysen et al. [12]: Dutch), the speaking rate, the intensity level, specified voice quality features (Gravano & Hirschberg [10]: American English), and an oscillatory speech rate (Couper-Kuhlen [13]: English; Auer et al. [14]: English, German, Italian; Beňuš et al. [15]: American English; Stivers et al. [16]: various languages; Wilson & Wilson [17]). Wilson and Wilson [17] suggested that the precise turn switches in conversation are possible because interlocutors align their cognitive cyclic patterns, that is, "an oscillatory function of readiness to initiate speech" (p. 962). They further stated that the speaker's syllable or speech rate is decisive for the frequency of the oscillation. Wagner et al. [18] likewise assumed that prosodic features of an utterance are rhythmically organized so that they follow a regular oscillation pattern. This pattern would simplify interlocutors' cognitive processing and interaction in communication and give them the opportunity to entrain their speech pattern (Inden et al. [19]: German). Thus, next speakers match their turn onset according to the entrained speech rate, achieving very efficient timing in turn-taking [13,14,17]. A third group of researchers argues in favor of several aspects – like semantics, syntax, prosody, rhythm, gesture, context, gaze and facial expression – as being altogether important for successful turn-taking processes (Ford & Thompson [20]: American English; Selting [21]: German).

This last approach implies that human communication is not only based on vocal but on visual clues as well. Interlocutors are able to communicate by gaze or gesture parallel to speech and thus indicate interest, repair or their aim to take the turn over. This is why the human turn-taking system is very efficient. In an electroencephalographic (EEG) study with German sentence stimuli, participants were instructed to press a button exactly at the moment an aurally presented sentence ended. A lateralized readiness potential related to the button press was observed 600 to 800 milliseconds before the end of the sentences, indicating that participants were aware of the sentence end 600 to 800ms before it actually ended [22]. Although only based on aural clues, this early anticipation leaves plenty of time for the listener to prepare their own turn and allows the turn transition to be smooth. Interestingly, in a follow-up study by Wesselmeier et al. [23] using the same sentence stimuli, the readiness potential was observed as early as 1.400ms before the end of a sentence but was disrupted for stimuli violating syntax or semantics. In those stimuli, the readiness potential occurred 900ms before the end of the sentence. The

findings of these studies suggest that upcoming speakers prepare for the end of their interlocutor's turn quite early, which supports an efficient turn-taking process [24]. In contrast to possible parallel means of communication in humans, complex vocalization in animals may only occur sequentially. Parallel vocalization would preclude successful communication. As a consequence, complex vocalization takes much more time than human communication and is thus pressed for time.

It is still under debate which aspects carry how much weight when it comes to end-of-turn anticipation. In the current study we aimed at finding out whether language, even if it is unknown to the listener, carries clues in the form of certain signals which indicate utterance-ends. Therefore, we compared the Anticipation Timing Accuracy (ATA) in six foreign languages, of which participants knew only English as their L2, and pure sinusoidal tones varying in their duration as a maximal contrast to natural linguistic stimuli. Anticipation processes are usually studied by the use of behavioral methods. The ATA is an indicator of such conscious behavioral processes and is a sign of how well subjects were able to anticipate the end of an utterance. The intention was to judge the importance of general acoustic natural speech-specific features other than content and grammar, like language universal suprasegmentals such as pitch movement, final lengthening, and other potential cues, for end-of-utterance anticipation.

2. Materials and Methods

Participants

37 students (18 women, 19 men) of Bielefeld University with German as their native language participated in the experiment. Written consent was obtained for publication of this study. Participants had a mean age of 23.7 years (±2.9) and were right-handed with a mean lateralization quotient of 88.1 (±13.3) according to the *Edinburgh Handedness Inventory* [25]. According to their own accounts, participants did not suffer from any auditory or motor restrictions or diseases which could have influenced results.

Stimuli

Stimuli were either spoken sentences (161 total) or pure 450 Hz sinusoidal tones of different durations (10 total). 23 spoken sentences were phrased in German and translated to English, Italian, Polish, Turkish, Arabic and Korean by native speakers of the respective languages, who were fluent in German. The recording of the spoken sentences was done in a sound-attenuated booth with the same native speakers who did the translation. Turkish and Arabic items were recorded with a male. German, English, Italian,

Polish, and Korean items with a female speaker.

All participants had good knowledge of English as their L2 (*M* = ten years of school education), two subjects had had marginal contact to Italian and one to Polish. All were unfamiliar with Arabic, Turkish and Korean. Except for German and English, all languages were judged as unknown foreign languages. The foreign language utterances were presented in order to test the influence of speech patterns and general language specific properties independent of semantics and syntax on end-of-utterance anticipation. The idea was that less participants could rely on semantic and syntactic content, the more they would have to anticipate the ends by use of the remaining language universal suprasegmental features or even more general but typical aspects of natural speech signals. If they are not trained through everyday conversation to do so, the anticipation performance of participants should be as bad for the linguistic but incomprehensible stimuli as they are for the maximally non-linguistic sinusoidal tones that do not contain any linguistic information at all. Other studies used low-pass-filtering or hummed speech to remove semantic content and syntactic structure [6,26]. Since the result of this technique makes the speech signal sound less like speech, we decided to use foreign language stimuli instead. These stimuli should have the same effect as hummed speech of filtered stimuli while remaining fully intact natural linguistic signals. Sentences had a mean length of 3591ms and were of a simple syntactic structure (e.g., *Books and articles about fossils are very interesting. Newspapers are a source of many different kinds of information.*). The ten sinusoidal tones were generated at 450Hz. They varied in length from 2600ms to 4400ms, had a mean length of 3300ms and thus matched the length of the sentences. In total, there were 171 stimuli.

We not only checked for ATA differences between the languages but also classified them as either Indo-European (IE: German, English, Italian, Polish) or non-Indo-European (non-IE: Arabic, Turkish, Korean) [27,28]. These two groups were then compared to each other and to the sinusoidal tones.

Procedure

All items were presented to all participants in pseudo-randomized order in a within-subject design. Following randomization, the order was checked manually so that no sentence type was followed by the same type (for instance, no English utterance followed an English utterance).

Prior to the experiment, participants were asked to fill out two questionnaires. The first asked for age, gender, course of study and motoric or sensory deficits (Appendix B). In the second, handedness was determined by the *Edinburgh Handedness Inventory* [25] (Appendix C). Participants practiced their task with six items (four German sentences, one foreign-language utterance and one

sinusoidal tone). Items were presented aurally via E-prime (Psychology Software Tools, vers. 2.0) on a PC (Windows 7). Subjects listened to the items with headphones. They then had the task of pushing a button on an external USB response box with an internal clock with the forefinger of their right hand at the exact moment the utterance ended [cf. 6]. If the button push occurred too early the utterance was stopped immediately and the next stimulus began after an ISI of 1000ms. Each item required a response in order for the subject to continue with the experiment. The external response box measured the ATA with an accuracy of about ± 2ms. The ATA was defined as the time span between the button push and the actual end of the utterance. After the procedure, language skills of participants in the foreign languages that came up in the experiment were documented in a separate questionnaire (Appendix D). As for the outcome of the experiment we posed the question whether it is possible to predict the ends of utterances in unknown foreign languages more accurate than the ends of sinusoidal tones.

3. Results

The statistical analysis of the resulting ATA was done via SPSS (IBM, vers. 20) under Mac OS X. First, descriptive statistics for each sentence type were calculated. Extreme values defined as values lower and higher than two standard deviations from the mean in each language amounted up to 3.1% of all valid responses and were excluded from the analysis. Further, four items were excluded from the analysis due to a low item-total correlation in the item reliability analysis (one German, Turkish, Arabic, and Korean item respectively). Subsequently, repeated measures ANOVAs were computed with ATA as the dependent and type of sentence as the independent variable with the factor levels *German*, *English*, *Italian*, *Polish*, *Turkish*, *Arabic*, *Korean* and *tones*. The Bonferroni multiple comparison test was done post-hoc.

Detailed results are available in Appendix A. There were several mentionable differences between the ATAs related to the different languages (Figure 1). Subjects anticipated the ends of German items better than of any other stimulus type in this category. Further, the ends of tones and of Turkish stimuli were anticipated equally worse than the ends of all other items. We tested the differences between the sentence types for their statistical significance. Since Mauchly's test was significant, the degrees of freedom were corrected by the Greenhouse-Geisser estimates of sphericity ($\varepsilon = 0.49$). Altogether, the factor levels had a highly significant impact on the ATA ($F(3.42, 119.52) = 100.27$, $p = .000$). The multiple comparison post-hoc test (Bonferroni) showed that almost all item types differed significantly from one another (for *p*-values see Appendix A). The differences of the ATA of tones and of German

(p= .000), Italian (p = .000), English (p = .000) and Arabic items (p = .000) were significant as were the differences between Turkish stimuli and German (p = .000), Italian (p= .000), English (p = .000), Polish (p = .001), Arabic (p= .000) and Korean items (p = .003) (see Figure 1 and Appendix A for further results).
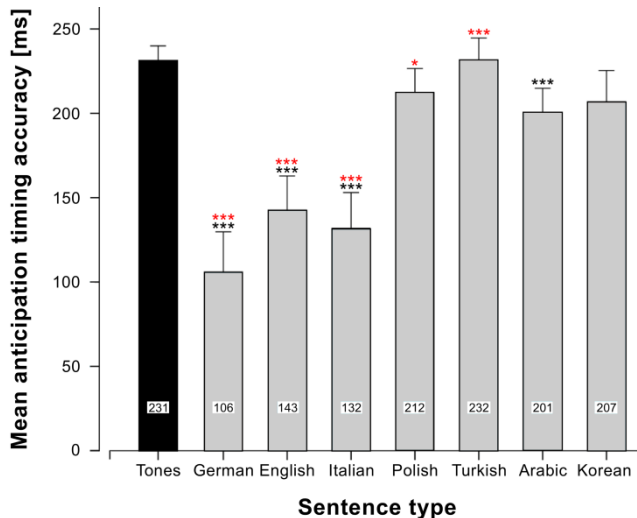
Significant differences to sinusoidal tones are marked by black asterisks, significant differences to Arabic items are marked by red asterisks (*** = $p \le$.001, * = $p \le$.05). Error bars represent 95% confidence interval.

Figure 1. Mean Anticipation Timing Accuracies (ATAs) for the different languages in comparison to sinusoidal tones

Categorizing languages to either IE or non-IE items and comparing them to the sinusoidal tones also revealed an overall highly significant effect (F(1.33, 46.65) = 98.35, p= .000). The post-hoc analysis showed a significant difference between the ATA on IE stimuli and on non-IE items (p = .000), on IE items and tones (p = .000) and on non-IE items and tones (p = .010). The ends of IE utterances were anticipated most accurately (Figure 2).
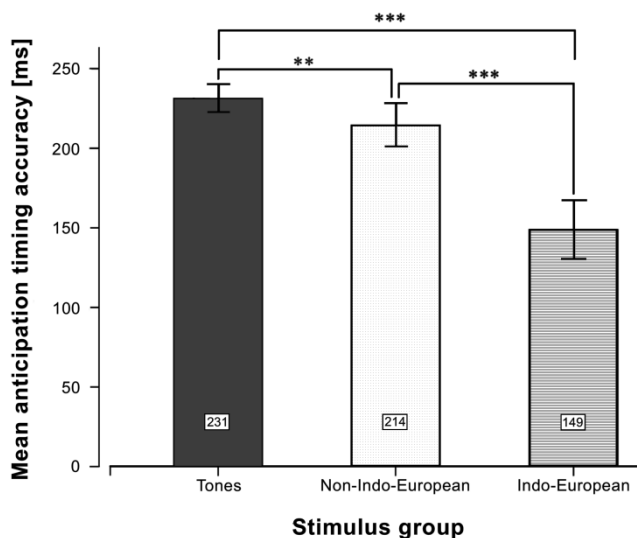
Significant differences are indicated by brackets (*** = $p \le$.001, ** = $p \le$.01). Error bars represent 95% confidence interval.

Figure 2. Mean Anticipation Timing Accuracies (ATA) of the grouped languages (Indo-European, Non-Indo-European) compared to sinusoidal

tones.

4. Discussion

The aim of the current experiment was to find out more about the influence of general language universals as part of a natural speech signal and speech-specific articulatory features on utterance-end anticipation. The question was whether anticipation performance would be better for end-of-utterance detection in unknown foreign languages than for ends of sinusoidal tones. Judging from the questionnaire (Appendix D), skills for all languages used in the experiment except German and English were either completely lacking or too little to influence the mean results in the current experiment.

Results showed that the ATAs of German, English, Italian, and Arabic items differed significantly from the ATA of tones. In contrast, there were no differences between the ATA of Polish, Turkish and Korean items and the ATA of tones. It thus seems to be the case that there is something familiar in Italian and Arabic speech stimuli which enables German native speakers to anticipate the ends of utterances in these languages. In German and English, participants could rely on semantics and syntactic structure, which becomes apparent in the low ATAs. Based on this result one can conclude that syntax and semantics have a great impact on anticipation performance. However, this is no explanation for the low ATAs in relation to Italian items and for the significant difference between Arabic stimuli and tones. In those languages, participants could have had access to suprasegmental characteristics of the utterances at most (and to single word semantics in the cases of Italian and Polish). However, since the ATA-differences between unknown languages and tones were predominantly not significant, it can be concluded that suprasegmental characteristics alone are not sufficient for an appropriate anticipation performance concerning utterance-ends if the languages are examined separately.

Grouping languages to IE utterances and non-IE items throws a somewhat different light on this outcome. Subjects anticipated the ends of IE utterances better than of utterances which are part of other language families. This might be due to the degree of similarity of single words or acoustic patterns in the related languages. Crucially, responses on sinusoidal tones were still significantly worse than on non-IE stimuli. Thus, if subjects did not have access to any semantic content or syntactic structure, as was the case for the non-IE items, they might have used other language universal linguistic features. Those features must have been of an acoustic, suprasegmental or stress related origin. Features that might have been relevant in this context are the last major accent and specific F_0-contours as identified in a number of corpus studies (Koiso et al. [29]: Japanese; Wells & Macfarlane [8]: English; Caspers [30]: Dutch; Heldner et al. [9]: Swedish). Wells and Macfarlane [8] compared the onsets of turn-competitive to non-competitive utterances by

recipients in a natural conversation in English. They found that the onsets of the competitive turns were usually placed right before the last major accent [8]. Therefore, the syllable carrying the last major accent is supposed to allow recipients the anticipation of the turn-end, given that it holds unique phonetic features, which were identified in a second experiment. Wells and Macfarlane [8] found a certain order of prosodic features – a step up and a drift down of pitch right after the last major accent. Additionally, the last major accent is pronounced louder and lengthened in contrast to accents which do not indicate a possible place for a speaker change. These characteristics make the last major accent distinct and set it apart from other accents making it possible for conversational partners to use it as an additional indicator to predict the end of a turn. In another corpus study conducted by Heldner et al. [9], results of earlier research concerning the relevance of F_0-patterns in turn-taking could be replicated for Swedish. They detected rising and falling pitch contours before the end of the turn. Before pauses that did not represent a speaker change, they found a rather flat pitch contour. As for non-IE languages, Koiso et al. [29] found suprasegmental cues, like the duration of the final phonemes, peak F_0 and peak energy, to be relevant for turn-taking in a Japanese corpus. All of these characteristics might also have been relevant for participants in the current study whenever they were not able to anticipate the utterance-end by use of semantic or syntactic criteria.

A further characteristic of language that contrasts it to the sinusoidal tones is a speech specific acoustic pattern, which might be used in end-of-utterance anticipation (Beňuš et al. [15]: American English). Auer et al. [14] define this pattern as the 'beat' of a language which constitutes a rhythmic isochrony which is then used by interlocutors to fit their turn onsets into the beat pattern. Based on this, oscillator models attempting to explain the efficient timing of interlocutors have been and are currently being developed. The idea behind such a model is that both speaker and listener are prepared to initiate speech at a frequency of an oscillation based on the speaker's syllable rate [17]. If this were true, participants might not have been able to detect the ends of sinusoidal tones in the current study due to a missing acoustic pattern or missing syllable boundaries. In the past, there have been attempts to classify languages according to their stress pattern. The outcome was that languages tend to fall in one of three categories: They may be judged as either stress-timed, syllable-timed or mora-timed [31]. This classification is based upon timing patterns in each language. Thus, stress-timed means that the time span between one stressed syllable and the next is always approximately of the same length and the same applies for syllable- and mora-timed languages. Whereas German, English and Arabic are stress-timed [31,32], the classification of the other languages used in the current study is disputed, or, rather, not trivial. Italian

seems to be either stress- or syllable-timed, according to the speaker's dialect. In due consideration of this influencing factor, our Italian stimuli should be classified as syllable-timed, since the native speaker came from Milan [31-33]. Polish is judged as definitely distinct from English [31,34] and Turkish as rather syllable-timed [35]. Korean is discussed as being either syllable-timed [36] or mora-timed [37]. Against the background of these attempts, one could expect that end-of-utterance anticipation is easier in languages with a stress pattern equal to that of the mother tongue. In the current study, this would apply to stress-timed languages. In fact, ATAs on stimuli in Arabic, which is one of the stress-timed languages judged as unknown in the current study, were significantly shorter than the ATAs on tones. ATAs on Polish, Turkish, and Korean utterances, which have a different stress pattern than German, did all not differ significantly from ATAs on tones. This could be an indicator of the acoustic pattern being relevant for the anticipation of an utterance-end, albeit the ATAs on Italian items seem not to fit the pattern. One might conclude that an inability to make use of either syntax and semantics or a well-known acoustic pattern leads to an inadequate anticipation performance. This indicates that the acoustic pattern probably is a relevant suprasegmental characteristic in the anticipation of utterance ends.

Our results support the view that conversational partners consider several aspects when anticipating the upcoming end of a turn. This view is also taken by Ford and Thompson [20] who analyzed turn changes from a corpus of two natural conversations in English and found that 71 % of all speaker changes could be contributed to syntactic and intonational completion in combination with pragmatic completion. The authors stated that a syntactic completion point is only interpreted as the end of a turn if it is further reinforced by the intonational contour and the pragmatic content of the utterance. Selting [21] agreed with this position and – based on her results from German conversations – claimed that syntactic and prosodic structures need to be considered equally in the analysis of end-of-turn projection. With the aim to improve a machine's detection of turn-ends, Edlund and Heldner [38] also found that intonation patterns, as detected in turns of a Swedish Map Task corpus, helped the machine to judge whether or not a silent pause actually was a turn-end. This supports the view that several aspects are relevant for turn-end-anticipation as well. Again, there are similar findings for Japanese [39].

5. Conclusions

As an answer to the research question posed in this study, language is special in the sense that it always provides information of some kind which supports efficient communication, even if only on a suprasegmental level.

We found a significant result stating that the ends of non-IE items could still be anticipated better than the ends of tones although subjects had no access to content or syntactic structure at all. In the current study, end-of-utterance anticipation was only based on aural clues. In combination with visual information, the outcome might have been even more prominent. It seems like there is some kind of intrinsic feeling or intuition for speech which enables listeners to "gain access" to information in languages they do not know. This shows that speech and language are more than their parts, more than syntax, semantics, and prosody. They carry additional information as a whole, by way of expression, melody and acoustic patterns, which we have not quiet understood so far.

Acknowledgements

Appendix A

Table 1. Results of the repeated measures ANOVAs and the Bonferroni post-hoc-test for the Anticipation Timing Accuracy (ATA) in the categories Languages and Language families; Non-significant differences are not included in the table

Languages					
rm ANOVA	**εGreenhouse-Geisser**	**Df**	**MS$_e$**	**F**	**p**
Mauchly's	.488	27			.000
Within-Subjects Effect (Greenhouse-Geisser corrected)		3.415	1747.285	100.269	.000
Bonferroni multiple comparison test					**p**
German vs. English					.000
German vs. Italian					.008
German vs. Polish					.000
German vs. Turkish					.000
German vs. Arabic					.000
German vs. Korean					.000
German vs. Tones					.000
English vs. Polish					.000
English vs. Turkish					.000
English vs. Arabic					.000
English vs. Korean					.000
English vs. Tones					.000
Italian vs. Polish					.000
Italian vs. Turkish					.000
Italian vs. Arabic					.000
Italian vs. Korean					.000
Italian vs. Tones					.000
Polish vs. Turkish					.001
Polish vs. Arabic					.038
Turkish vs. Arabic					.000
Turkish vs. Korean					.003
Arabic vs. Tones					.000
Language families					
rm ANOVA	**εGreenhouse-Geisser**	**Df**	**MS$_e$**	**F**	**p**
Mauchly's	.666	2			.000
Within-Subjects Effects (Greenhouse-Geisser corrected)		1.333	1040.382	98.347	.000
Bonferroni multiple comparison test					**p**
IE vs. Non-IE					.000
IE vs. Tones					.000
Non-IE vs. Tones					.010

Appendix B

Questionnaire 1. The background questionnaire asking participants for age, gender, course of study and motoric or sensory deficits among other things.

VP: _____　　Datum: _____　　Uhrzeit: _____

Alter: _____　　Geschlecht: _____　　Gruppe: _____

Studiengang: _____

Tragen Sie eine Brille oder Kontaktlinsen?　　ja☐　nein☐
Tragen Sie ein Hörgerät?　　ja☐　nein☐

Ist Deutsch Ihre Muttersprache?　　ja☐　nein☐
Sprechen Sie weitere Muttersprachen?　　ja☐　nein☐
Wenn ja, welche? _____

Sind Sie aus irgendwelchen Gründen in den Bewegungen des rechten Arms und der rechten
Hand eingeschränkt?　　ja☐　nein☐
Wenn ja, bitte erläutern Sie:

_____
_____

Sonstige Bemerkungen:

_____
_____

Appendix C

Questionnaire 2. The Edinburgh Handedness Inventory as introduced by Oldfield [25] and as used in the current study.

112   R. C. OLDFIELD

APPENDIX II

Medical Research Council Speech & Communication Unit

EDINBURGH HANDEDNESS INVENTORY

Surname.. Given Names..

Date of Birth.. Sex..

Please indicate your preferences in the use of hands in the following activities *by putting* + *in the appropriate column.* Where the preference is so strong that you would never try to use the other hand unless absolutely forced to, *put* + +. If in any case you are really indifferent *put* + *in both columns.*

Some of the activities require both hands. In these cases the part of the task, or object, for which hand preference is wanted is indicated in brackets.

Please try to answer all the questions, and only leave a blank if you have no experience at all of the object or task.

		LEFT	RIGHT
1	Writing		
2	Drawing		
3.	Throwing		
4	Scissors		
5	Toothbrush		
6	Knife (without fork)		
7	Spoon		
8	Broom (upper hand)		
9	Striking Match (match)		
10	Opening box (lid)		
i	Which foot do you prefer to kick with?		
ii	Which eye do you use when using only one?		

L.Q. [   ]   Leave these spaces blank   DECILE [   ]

MARCH 1970

Appendix D

Questionnaire 3. The questionnaire asking subjects to indicate their language skills in the foreign languages that came up in the experiment

VP: _____

Haben Sie eine der folgenden Sprachen als weitere Muttersprache erlernt?

Sollten Sie eine der folgenden Sprachen nicht muttersprachlich erworben haben, geben Sie bitte an, wie viele Jahre Sie diese Sprache gelernt haben.

Englisch

Muttersprache ☐   _____ Jahre gelernt   keine Kenntnisse ☐

Italienisch

Muttersprache ☐   _____ Jahre gelernt   keine Kenntnisse ☐

Polnisch

Muttersprache ☐   _____ Jahre gelernt   keine Kenntnisse ☐

Türkisch

Muttersprache ☐   _____ Jahre gelernt   keine Kenntnisse ☐

Arabisch

Muttersprache ☐   _____ Jahre gelernt   keine Kenntnisse ☐

Koreanisch

Muttersprache ☐   _____ Jahre gelernt   keine Kenntnisse ☐

REFERENCES

[1] E. A. Schegloff. Turn organization: One intersection of grammar and interaction. E. Ochs, E. A. Schegloff, Thompson (Eds.), Interaction and grammar, Cambridge University Press, Cambridge, 52-133, 1996.

[2] K. Brinker, S. F. Sager. Linguistische Gesprächsanalyse. Eine Einführung, Schmidt, Berlin, 2010.

[3] S. Yoshida, K. Okanoya. Animal cognition evolution of turn-taking: A bio-cognitive perspective, Cognitive Studies, 12, 153-165, 2005.

[4] H. Sacks, E. A. Schegloff, G. Jefferson. A simplest systematics for the organization of turn-taking for conversation, Language, 50, 696-735, 1974.

[5] G. W. Beattie. The regulation of speaker turns in face-to-face conversation. Some implications for conversation in sound-only communication channels, Semiotica, 34, 55-70, 1981.

[6] J. P. de Ruiter, H. Mitterer, N. J. Enfield. Projecting the end of a speaker's turn. A cognitive cornerstone of conversation, Language, 82, 515-535, 2006.

[7] L. Magyari, M. C. M. Bastiaansen, J. P. de Ruiter, S. C. Levinson. Neuronal correlates of anticipation related to turn-taking in conversation, AMLaP 2011: Architectures and Mechanisms for Language Processing, 75-76, 2011.

[8] B. Wells, S. Macfarlane. Prosody as an interactional resource: turn-projection and overlap, Language and Speech, 41, 265-294, 1998.

[9] M. Heldner, J. Edlund, K. Laskowski, A. Pelcé. Prosodic features in the vicinity of silences and overlaps, Nordic Prosody, 95-106, 2009.

[10] A. Gravano, J. Hirschberg. Turn-yielding cues in task-oriented dialogue, SIGDIAL 2009: the 10th Annual Meeting of the Special Interest Group in Discourse and Dialogue, 253-261, 2009.

[11] E. Dombrowski, O. Niebuhr. Acoustic patterns and communicative functions of phrase-final F_0 rises in German: Activating and restricting contours, Phonetica, 62, 176-195, 2005.

[12] P. Barkhuysen, E. Krahmer, M. Swerts. The interplay between the auditory and visual modality for end-of-utterance detection, The Journal of the Acoustical Society of America, 123, 354-365, 2008.

[13] E. Couper-Kuhlen. English speech rhythm: Form and function in everyday verbal interaction, Benjamins, Amsterdam, 1993.

[14] P. Auer, E. Couper-Kuhlen, F. Müller. Language in Time, Oxford University Press, Oxford, USA, 1999.

[15] Š. Beňuš, A. Gravano, J. Hirschberg. Pragmatic aspects of temporal accommodation in turn-taking, Journal of Pragmatics, 43, 3001-3027, 2011.

[16] T. Stivers, N. J. Enfield, P. Brown, C. Englert, M. Hayashi, T. Heinemann, G. Hoymann, F. Rossano, J. P. de Ruiter, K.-E. Yoon, S. C. Levinson. Universals and cultural variation in turn-taking in conversation, Proceedings of the National Academy of Science, 106, 10587-10592, 2009.

[17] M. Wilson, T. P. Wilson. An oscillator model of the timing of turn-taking. Psychonomic Bulletin and Review, 12, 957-968, 2005.

[18] P. Wagner, S. Malisz, B. Inden, I. Wachsmuth. Interaction phonology – a temporal coordination component enabling representational alignment within a model of communication, I. Wachsmuth, J. P. de Ruiter, P. Jaecks, S. Kopp (Eds.), Advances in interaction studies. Alignment in communication, Benjamins, Amsterdam, 109-132, 2013.

[19] B. Inden, S. Malisz, P. Wagner, I. Wachsmuth. Timing and entrainment of multimodal backchanneling behavior for an embodied conversational agent, 15th International Conference on Multimodal Interaction, 181-188, 2013.

[20] C. E. Ford, S. A. Thompson. Interactional units in conversation. Syntactic, intonational, and pragmatic resources for the management of turns, E. Ochs, E. A. Schegloff, S. A. Thompson (Eds.), Interaction and grammar, New York, Cambridge University Press, 134-184, 1996.

[21] M. Selting. On the interplay of syntax and prosody in the constitution of turn-constructional units and turns in conversation, Pragmatics, 6, 357-388, 1996.

[22] S. Jansen, H. Wesselmeier, J. P. de Ruiter, H. M. Müller. Using the readiness potential of button-press and verbal response within spoken language processing, Journal of Neuroscience Methods, 232, 24-29, 2014.

[23] H. Wesselmeier, S. Jansen, H. M. Müller. Influences of semantic and syntactic incongruence on readiness potential in turn-end anticipation, Frontiers in Human Neuroscience, 8, 296, 2014.

[24] S. Bögels, S. C. Levinson. The brain behind the response: Insights into turn-taking in conversation from neuroimaging, Research on Language and Social Interaction, 50, 71-89, 2017.

[25] R. C. Oldfield. The assessment and analysis of handedness: The Edinburgh Inventory, Neuropsychologia, 9, 97-113, 1971.

[26] A. K. Ischebeck, A. D. Friederici, K. Alter. Processing prosodic boundaries in natural and hummed speech. An fMRI study, Cerebral Cortex, 18, 541-552, 2008.

[27] B. Comrie (Ed.). The world's major languages, Oxford University Press, Oxford, USA, 1987.

[28] M. Ruhlen. A guide to the world's languages, Stanford University Press, Stanford, USA, 1987.

[29] H. Koiso, Y. Horiuchi, S. Tutiya, A. Ichikawa, Y. Den. An analysis of turn-taking and backchannels based on prosodic and syntactic features in Japanese map task dialogs, Language and Speech, 41, 295-321, 1998.

[30] J. Caspers. Local speech melody as a limiting factor in the turn-taking system in Dutch, Journal of Phonetics, 31, 251-276, 2003.

[31] F. Ramus, M. Nespor, J. Mehler. Correlates of linguistic rhythm in the speech signal, Cognition, 73, 265-292, 1999.

[32] M. Russo, W. J. Barry. Measuring rhythm. A quantified analysis of Southern Italian dialects stress time parameters, Language Design, 2, 315-322, 2008.

[33] B. Braun, S. Geiselmann. Italian in the no-man's land between stress-timing and syllable-timing? Speakers are more stress-timed than listeners, 12th Annual Conference of the International Speech Communication Association, 2697-2700, 2011.

[34] F. Ramus, E. Dupoux, J. Mehler. The psychological reality of rhythm classes: Perceptual studies, 15th International Congress of Phonetic Sciences, 337-342, 2003.

[35] P. Mairano, A. Romano. Testing the perception of speech rhythm on natural and artificial stimuli, 6th International Conference on Speech Prosody, 482-485, 2012.

[36] P. Mok, S. I. Lee. Korean speech rhythm using rhythmic measures, 18th International Congress of Linguistics, 2008.

[37] Z. S. Bond, V. Stockmal. Distinguishing samples of spoken Korean from rhythmic and regional competitors, Language Sciences, 24, 175-185, 2002.

[38] J. Edlund, M. Heldner. Exploring prosody in interaction control, Phonetica, 62, 215-226, 2005.

[39] C. Nakamura, M. Arai, R. Mazuka. Immediate use of prosody and context in predicting a syntactic structure, Cognition, 125, 317-323, 2012.