Prosodic Universals in Discourse Particles

Tillmann Pistor

Forschungszentrum Deutscher Sprachatlas
Philipps-University of Marburg, Germany
tillmann.pistor@staff.uni-marburg.de

Abstract

Recent prosodic research has shown that discourse particles such as *hm* or *äh* offer an optimal basis for exploring the functions of prosody in general, since these forms are simply free of semantic and segmental information. This paper outlines the results of a comparative study focusing on prosodic structures as well as their functions occurring in discourse particles in five different and non-related language families. Results of testing perception and production show that there are at least four prosodic units which are similar in their phonetic form as well as in their phonological function in all five language families tested and may thus be regarded as potential language universals.

Index Terms: prosody, intonation, language universals, discourse pragmatics, emotion

1. Introduction

Many studies dealing with the description of the prosody of a language leave numerous gaps in the description itself, which primarily stems from either focusing on pitch only (often having a simple setting of only two tonal levels), or from not dealing with the delineation of units investigated in the speech-flow. Investigating isolated items only, this study does not have to deal with this problem of delineation in the first place. Furthermore, a model of describing the prosody of a language was applied, which sheds light on all prosodically relevant features including pitch, intensity and signal duration. The entirety of this model cannot be dealt with in this paper, but has been described elsewhere (cf. [8], [13], [14]). However, the most essential features of the model will be briefly described in the following paragraph.

Prosodic units investigated in this study are according to the applied model *local intonational patterns*, with the syllable as their prosodic domain and the (auditory relevant) pitch contour as the primarily relevant prosodic feature, which can be based on a complex of acoustic and articulatory parameters (cf. [8]). The relation of phonetic form and function can either be discrete (a specific constellation of features marking a specific function, a change in these features thus means a change in the function) or continuous (a more or less of a feature or features gradually modifies the function).

Based on a pilot study investigating the prosody in discourse particles in German (cf. [9], [14]), the functions of these prosodic units may primarily (a) discourse-pragmatically organize communication, and (b) communicate emotionally the attitude of the speaker. In the model, particles having the function of (a) can be differentiated between working progressively (directly influencing the following communication structure) and working reactively (referring to the preceding turn).

In this study, the hypothesis was advanced that the functions of prosodic units stated above have cross-linguistic validity when the units occur in discourse particles. The following three research questions guided this empirical study: 1. Are there prosodic universals in discourse particles, and if so, how many are there? 2. What are their prototypical forms and are there phonetic variants of those forms? 3. To which functional categories can these forms be allocated and are these functions universal as well?

2. Methods and analyses

The investigation of potentially universal prosodic units was conducted in two main studies. The aim of study A was to gather spoken data and by doing so, create inventories of discourse particles from five different language families and then comparatively analyze the phonetic forms and phonological functions. Study B aimed to validate the functional categories, which the units resulting from study A were allocated to. The methods from both studies will be presented in the following paragraphs.

2.1. Study A

Presenting evidence for potential universals demands data from speakers of languages which differ maximally, and have a non-relational status in terms of general language typology. Therefore, two native speakers from each of the following languages were consulted: German, Chinese (Mandarin), Gomala (a dialect in western Cameroun), Arab and Korean. This range of languages covers intonational languages as well as tone languages, whereby none of these are related typologically (cf. [2], [10]). The speakers were chosen on the basis of language use and skills, in which the use of the targeted native language must outweigh any other languages spoken by the candidate. Speakers were to observe their use of discourse particles in interactions with other native speakers for a period of at least two weeks. These items were then recorded in various contexts (e.g. isolated, sentence, fictional dialogue) in which the speakers were asked to give a brief subjective description of the function (e.g. by direct speech, description of a situation) as well. All produced items were then phonetically measured using the speech analysis program *praat* (cf. [1] with regard to all three relevant acoustic parameters (pitch, intensity, duration) and compared to each other, whereby the segmental basis was also considered (e.g. vowel vs. consonant, glottal stop vs. aspiration). Absolute F0-values were converted to relative semitones to enable a direct comparison of female and male spoken data. The following formula was applied (cf. [7]):

$$A_{(HT)} = 12 \times log_2\left(\frac{f1}{f2}\right) \qquad (1)$$

Using the similarity from a) the measurements of the parameters above and b) the functions given by the speakers as criteria, final prototypes were identified. Based on a theory of cognitive psychology (cf. [12]), a prototype is defined here as the most central member of a category, which joins together all features that occur with the highest average in all other members of the same category. Thus, recurring phonetic forms from all speakers were filtered out of a total of 61 items and audibly presented to a total of 71 listeners from various native language backgrounds. Listeners were not given any further information or the communicative contexts of the isolated signals. Initially, all listeners were to evaluate the signals on whether they exist in their own individual mother tongue in questionnaires, and were secondly asked to write down in direct speech, what the speaker could have had said instead of the signal that they heard (*verbalization task*). Such statements allow direct access to the concrete communicative function(s) of the items, as well as insight into the possible linkage of functions (discrete vs. continuous, see above) as well.

2.2. Study B

This study aimed to validate the functional categories, to which the units resulting from study A were allocated. Therefore, all listener statements (henceforth called *production contexts*) from the verbalization task in study A were, if showing obvious synonymy, subsumed under the category of items, to which they were produced, and subsequently arranged to their frequency. The three most frequent production contexts for each category were then printed on playing cards, adapted from the party game "Tabu", in which player A has to explain the content of his card to player B without using specific terms listed on the card. In my version of the game, player A was not allowed to open the mouth while explaining the production context, which showed e.g. a sentence or exclamation on the card (e.g. "What did you say?", "This is/smells good!"). The visual channel was excluded. The set of cards contained 15 relevant production contexts and 5 distractions. They were presented to the player in a randomized order.

For this study, 16 participants were acquired, corresponding to a total of 8 groups of 2 players each. All participants only spoke German as a native language. A card, and thus a production context, was successful, when player B was able to recognize player A's intention within 30 seconds of explanation time per card. Thus, the number of production contexts was reduced, which made it possible to narrow down the functions.

The items that player A produced were additionally phonetically measured and compared to the prototypical forms of study A. Results show, that these forms are strikingly similar, as expected.

3. Results

The observation of having both equal forms and functions in all 5 language families tested leads to the assumption, that these prosodic units are potential universals. Furthermore, the combined results from both studies show four functional prosodic units on discourse particles, based on an unstrained vowel or consonant with optional initial glottal stop or aspiration. These units manifest themselves as local intonational patterns, as described above, showing both discrete and continuous relations of form and function. Descriptions of the proto-

typical forms and functions of the units will be given in the following paragraphs.

3.1. Reaction

The first potentially universal prosodic unit was called *reaction* and shows a short, locally rising intonational pattern realized in an average register of pitch. F0 rises either convexly or in a straight line, extending over an average duration of 0.23 seconds. Its prototypical verbalization is "what/what did you say". The universal function of this discrete prosodic unit is the organization of communication. It is communicated reactively by the speaker referring to the preceding turn, calling upon the recipient to react. The potential universality of this unit has also been examined and confirmed elsewhere (apart from two exceptions, cf. [3]). The relevant prosodic feature is the auditory relevant pitch contour.
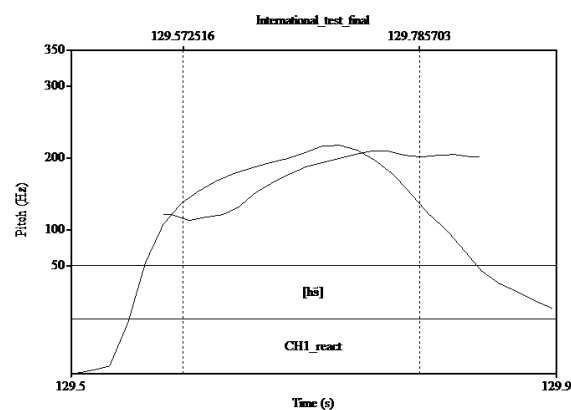

Figure 1: Acoustic analysis of reaction-signal

Figure 1 shows an example of the acoustic analysis from the Chinese inventory. The illustration depicts F0 in its rising contour, and intensity as the rising-falling contour. In all figures, the measured intensity values were not depicted, since they have no influence on the functions of the units. The dotted lines show the relevant area of measurement. The speaker realized this item upon [hɜ̃] – an unstressed nasalized vowel with preceding aspirational onset.

3.2. Turn holding

The second universally occurring unit was called *turn holding*. The pitch of this intonational pattern remains at a relatively equal level or declines gently for the most part of the long average duration of 0.79 seconds. The prototypical verbalization is "let me think/I'm not sure". This discrete prosodic unit also serves the purpose of organizing communication and signalizes progressively that communication on the part of the uttering speaker requires continuation. The relevant prosodic features are the auditory relevant pitch contour and duration.
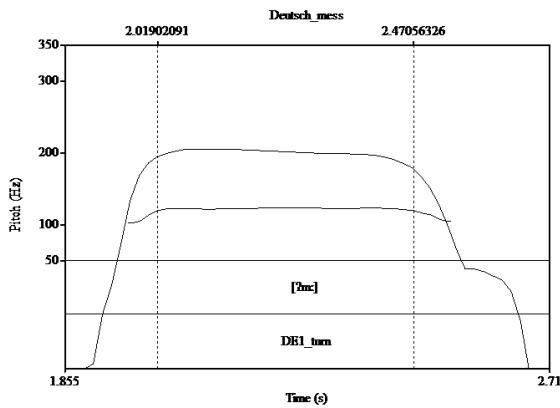
Figure 2: Acoustic analysis of holding turn

In figure 2, the smaller line around 100 Hz depicts the F0 contour. In this contour, the areas outside the dotted lines can be regarded as signal redundancy (i.e. having no effects on the perception or the interpretation of the signals), marking the area within the dotted lines as the perceptually relevant part of the signals in this category. The German speaker realized this item upon [ʔmː] – a long bilabial consonant with preceding glottal stop.

3.3. Responding

The third potential universal is a category of units, containing two short, falling intonational patterns which appear similar at first sight, but show distinct forms and functions according to the analyses. The units were called *closure* and *taking cognizance* and were subsumed under the category *responding*, since both units reactively refer to the preceding turn, function as closing signals to the turn itself or its content, and ascribe this category a communication organizing function. Both signals show an average duration of 0.20 seconds. They will be depicted in the figure below and subsequently described in two short paragraphs.
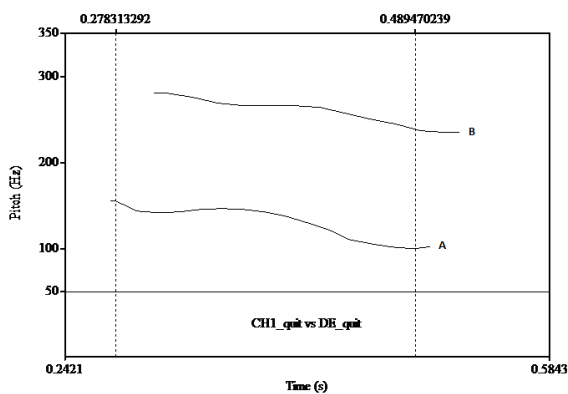

Figure 3: F0 comparison of closure (A) and taking cognizance (B)

3.3.1 Closure

Discrete prosodic units holding the function of closure, prototypically show a strong decline of F0, either concavely (as depicted above in line A; Chinese native speaker) or in a straight line. These items are realized in a low register of pitch for the

individual speaker. An initial glottal stop is optional. Its prototypical verbalization is "okay/done". The unit again serves the purpose of organizing communication. The unit is intuitively produced by all speakers as a discourse particle, but is preferably perceived by listeners as one-word-utterances. The relevant prosodic features are the auditory relevant pitch contour and the register of pitch, which are distinguishing units of this category to those described in the next paragraph.

3.3.2 Taking cognizance

Prosodic units of this category are ascribed a continuous relation of form and function. In this relation, the speaker's choice of pitch register, in which the signals are being realized, and the use of initial glottal stops cause either a connotation of negative (low register, glottal stop) or neutral (average/high register, no glottal stop) speaker attitude towards the main function. At the same time, these units are being realized with a weaker decline of F0 (as depicted above in line B; German native speaker), compared to signals of the function *closure*. The prototypical verbalization is "well/don't care", and the relevant prosodic features are again the auditory relevant pitch contour and the register of pitch. These units serve both to give a statement about the content, and close the preceding turn as well.

3.4. Positive evaluation

The fourth and last potentially universal prosodic unit named *positive evaluation* describes a long, complexly falling-rising-falling F0 contour. Here, the rising-falling part and especially the final decline of F0 claim most of the average duration of 0.79 seconds. The prototypical verbalization of signals in this category is "that is/tastes/smells good/delicious". The universal function of this prosodic unit is to communicate a positive speaker attitude, making this unit a genuinely emotional prosodic unit, as described in a previous study (cf. [8]). The relevant prosodic features are the auditory relevant pitch contour and the signal duration.
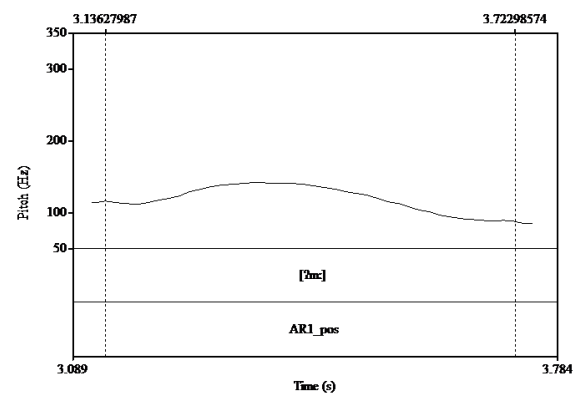

Figure 4: F0 contour of positive valuation

The figure above shows the signal allocated to *positive valuation* produced by a native speaker of Arabic. The short initial decline of F0 is optional in these items; a phonetic variant of this form only shows the rising-falling contour.

4. Conclusions

In this study, four prosodic units with discrete and one with continuous relation of form and function could be shown as corresponding regarding phonetic forms or their phonetic variants and phonological functions in five different language families. The units are manifested in local intonational patterns on discourse particles. In accordance with the terminology of language typology and universals (cf. [4], [5], [6], [7]), these findings confirm the hypothesis that these units can be regarded as potential prosodic-phonological universals of language with heuristic character. Furthermore, according to the units investigated, the findings of their universally occurring functions allow the statement that prosody in general holds the functions of primarily discourse-pragmatically organizing communication, and secondarily emotionally communicating the attitude of the speaker.

The order of these functions may also be regarded as a hierarchic structure, since units that organize communication occurred most frequently in all five language families investigated. Units communicating speaker attitudes and those serving as statements about content occur equally frequent, whereas for the percipient, discourse organizing and emotional units are easier to identify without any given syntagmatic, communicative or situational context, as the content units.

Questions of completeness of this heuristic inventory of prosodic universals and the possibility of transferring the outlined functions onto larger syntagmatic units such as words or phrases and finally their interaction with other language units, composing themselves to intonationally and syntactically complex utterances remain tasks for future studies for the moment.

5. References

[1] Boersma, Peter/Weenink, David: *Praat. Doing phonetics by computer*. <http://www.fon.hum.uva.nl/praat/>.

[2] Campbell, George L./King, Gareth (2013): *Compendium of the World's languages. Volume I: Abaza to Kyrgyz*. London/New York: Routledge.

[3] Enfield, Nick et al. (2013): Huh? What? A first survey in twenty-one languages. In: HAYASHI, M./RAYMOND, G./SIDNELL, J. (Hrsg.): Conversational repair and human understanding. New York: Cambridge University Press, S. 343-380.

[4] Greenberg, Joseph H. (1966): *Language Universals. With special reference to feature hierarchies*. The Hague (u.a.): Mouton (Janua linguarum: Series minor 59).

[5] Greenberg, Joseph H. (1974): *Language typology. A historical and analytic overview*. The Hague (u.a.): Mouton (Janua linguarum: Series minor 184).

[6] Greenberg, Joseph H. (1978): *Universals of Human Language 2: Phonology*. Stanford: Stanford University Press.

[7] Haase, Martin (2001): *Sprachtypologie und Universalien bei Joseph Greenberg*. In: Haspelmath, Martin/König, Ekkehard/Oesterreicher, Wulf/Raible, Wolfgang (eds.): *Sprachtypologie und sprachliche Universalien*. Handbücher zur Sprach- und Kommunikationswissenschaft 20. – Berlin/New York: de Gruyter 2001: § 23, I, S. 280 – 283.

[8] Kehrein, Roland (2002): *Prosodie und Emotionen*. Tübingen: Niemeyer (Reihe Germanistische Linguistik 231).

[9] Kehrein, Roland/Rabanus, Stefan (2001): *Ein Modell zur funktionalen Beschreibung von Diskurspartikeln*. In: Schmidt, Jürgen Erich [Hrsg.] (2001): *Neue Wege der Intonationsforschung*. Hildesheim: Olms (Reihe Germanistische Linguistik 157 – 158, S. 33 – 50).

[10] Pereltsvaig, Asya (2012): *Languages of the World: an Introduction*. Cambridge: Cambridge University Press.

[11] Rabanus, Stefan (2001*): Intonatorische Verfahren im Deutschen und Italienischen*. Linguistische Arbeiten 439. Tübingen: Niemeyer.

[12] Rosch, Eleanor (1973): Natural Categories. In: Cognitive Psychology 4, 328 – 350.

[13] Schmidt, Jürgen Erich (1986): *Die mittelfränkischen Tonakzente (Rheinische Akzentuierung)*. Stuttgart: Steiner (Mainzer Studien zur Sprach- und Volksforschung 8).

[14] Schmidt, Jürgen Erich (2001): *Bausteine der Intonation?*. In: Schmidt, Jürgen Erich [Hrsg.] (2001): *Neue Wege der Intonationsforschung*. Hildesheim: Olms. (Reihe: Germanistische Linguistik 157 – 158, S. 9 – 32).