

Two contours, two meanings: the intonation of *jaja* in German phone conversations

Andrea Golato* & Zsuzsanna Fagyal[#]

Departments of Germanic Languages and Literatures* and French[#]
University of Illinois at Urbana-Champaign
{golato; zsfagyal}@uiuc.edu

Abstract

This paper shows that *jaja* ‘yes yes’ sequences in German conversations carry two distinct interactional meanings cued by their intonation and sequential placement. Combined Conversation Analytic (CA) and Intonation Phonological analyses indicate that *jaja* tokens uttered with H* L-% intonation (following GToBI [10]) convey that the previous speaker has persisted too long in a specific course of (verbal) action which should therefore be stopped. By contrast, *jaja* tokens with L+H* L-% intonation are used in situations of fractured intersubjectivity, i.e., immediately after speakers misalign: with the *jaja* turn, its speaker treats the action/content of the previous speaker’s utterance as either unwarranted or self-evident. Speaking rate and regional dialectal differences notwithstanding, the two types of contour show significantly different peak alignment, and correspond to two distinct ‘peak accent’ nuclear contours.

1. Introduction

CA studies increasingly include prosodic features in analyses of talk-in-interaction [4], [5], [15]. Although typically studying intonational variation in more controlled contexts (reading, map tasks, guided interviews), phonologists have also shown interest in the analysis of natural discourse (see e.g. [12]). However, despite calls for using CA methods in phonetically-grounded intonational analyses, research in these two fields has so far progressed independently. In this paper, we combine CA and intonation phonology to inform both fields on the tonal-interactional meaning of intonation contours, and to “renew the connection of the analysis with the behaviour of everyday speakers” [12: 178].

We focus on a specific response token, the doublet *jaja*, in naturally occurring German telephone conversations. For intonation phonological analysis, such tokens present the advantage of occurring frequently, carrying invariable lexico-semantic meaning, and being composed of two identical syllables with all-voiced segments on which tone alignment is easy to track. Also, segment-induced f₀ variations, if not absent, are comparable between the two syllables. Work on response tokens in CA has yielded systematic functional differences between individual tokens [8], [13], [17], [20]. Previous research has shown that such tokens are a means by which interlocutors routinely provide the current speaker with information as to how their talk was understood and receipted [8] and how the recipient aligns with the current utterance [8], [20].

In this paper we argue that while single *ja* ‘yes’ tokens are typically analyzed as acknowledgement tokens, confirmation markers, or continuers (e.g., [3], [9]), a doubled *ja*, either produced as a H* L-% (^*jaja*) or a L+H* L-% (*ja*^*ja*)

contour cannot simply be considered a more intense version of the same action. In fact, these two forms, we argue, systematically accomplish separate interactional goals. By uttering *jaja* with what is termed a ‘peak accent’ (H* L-%) nuclear contour in GToBI [10: 65], the speaker conveys that the previous speaker has persisted too long in a specific course of action and that therefore this action should be stopped. By contrast, when speaker A produces a ‘rise from low up to peak accent’ (L+H* L-%) nuclear contour on the *jaja* token in response to an utterance by speaker B, he/she indicates that speaker B is misaligned with an earlier utterance by speaker A.

2. Method

Data were taken from a corpus of 9 hours of naturally occurring, audiotaped telephone conversations in German between close friends and family. Participants were only recorded during activities that they would normally engage in with each other. Speakers were from middle or upper-middle class and came from four different dialect areas in Germany. The recordings were transcribed using the CA transcription notation developed by Jefferson [2]. A total of 54 *jaja* tokens were identified perceptually, and analyzed using CA.

Independently from the CA examination, the 54 tokens were then isolated from their original context for acoustic-phonetic examination. 11 isolated tokens were excluded as they showed overlap with the other speaker. The remaining 43 *jaja* tokens were analyzed using PRAAT. The tokens were first segmented into syllables. The boundary between the vowel /a/ and the following palatal approximant /j/ was placed at the onset of the F2 transition between /a/ (low F2) and /j/ (high F2); the segmented syllable was then auditorily verified. In eight cases, when spectrographic information was unclear, the boundary was determined based on auditory information alone. This procedure was performed using the same procedure three times in three different sessions. Coding consistency showed a high intersession correlation ($r^2 \geq .89$), with an average difference of 8.3 ms between corresponding syllable boundaries.

The lowest f₀ values or ‘valleys’ (possible L tonal targets) and the highest f₀ values or ‘peaks’ (possible H tonal targets) in both syllables of each token were identified using the built-in functions ‘get minimum/maximum pitch’ in PRAAT. A cut-off point of 9Hz was applied, below which f₀ variations were attributed to noise or intrinsic pitch differences between the open vowel and the palatal approximant. Creaky voice and background noise led to several false positives, which were manually checked and discarded when necessary.

3. Results: Two contours

Only one major f₀ ‘peak’—not caused by noise, creaky voice, or segmental variations—was found in each *jaja* token. These

points were taken to be anchor points of H tones, and plotted with respect to the onset of the second *ja* syllable for all tokens. These alignment patterns were then contrasted with conclusions drawn from CA.

Figure 1. Peak alignment with respect to the onset of the second *ja* in two types of *jaja* token.

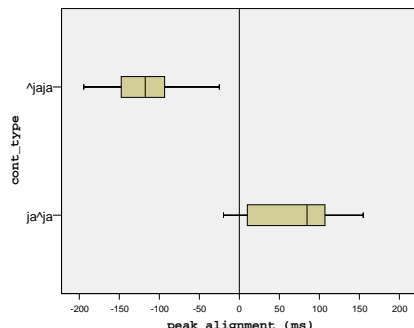


Figure 1 shows patterns of peak alignment in the 43 *jaja* tokens, matched to the categories established independently within the CA framework. The contours exhibit an almost categorical split between a group of contours (henceforth, *^jaja*) in which the f0 peak is anchored with the first syllable ($n=21$), and a group (henceforth, *ja^ja*) where the peak is associated with the second syllable ($n=22$). A one-way ANOVA comparing the two distributions shows a highly significant statistical difference ($F(1, 41) = 131.44$; $p < 0.001$). The same pattern obtains when peak alignment is plotted as % of the total duration of the word. While the peak in the *^jaja* contours is invariably located in the first syllable, the *ja^ja* group shows a much larger within-group variation. Peak alignment is highly variable within the second syllable; in fact, in four out of twenty-two cases the peak is realized at or near the offset of the first syllable (see section 5 for more detail).

Patterns of ‘valley’ (possible L tone) alignment show a similar, near-categorical, division. Most contours in the *^jaja* group—even though some show a slight final rise—tend to be falling towards the lowest f0 target point placed at the end of the second syllable. In most contours in the *ja^ja* group, on the other hand, the lowest f0 point occurs somewhere in the first syllable. Thus, acoustic-phonetic data indicate two distinct types of contour.

4. Results: Two meanings

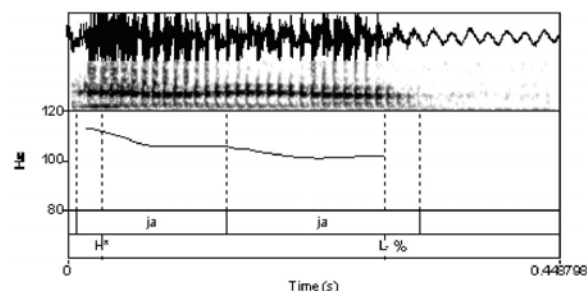
The data segment below shows an instance of a *^jaja* token in context (see line 6 and Figure 2). In this data segment, M is talking to his grandmother O who is reporting on a story she has read in the paper:

- 1 O: *und da steht heut die woch inner zeitung da*
 un there is today this week in the newspaper there
 2 O: *is:- hat eine f- du kennst doch diese stände aufm*
 is:- has a f- you know PRT those booths at+the
 3 O: *ma:kt, mit dem fleischerwagen.*
 ma:ket, with the butcher’s cart.
 4 M: *ja.*
 yes.
 5 O: *>ä< s:o am markttag gä*
 >uh< l:ike on market day right
 6 M: => *jaja. untn am do:m,*
 yes yes. down at the cathedral.
 7 O: *un’d ja:.. un da hat eine: hat eine familie*

- an’t^h ye:s. un there a family bought
 8 O: *fleischsalat gekauft, und is’ nach hause gegangen,*
 diced meat salad, and went home,
 9 O: *wolln den fleischsalat essen .hh und da is’ ein*
 wanted to eat the diced meat salad, .hhh and there’s
 10 O: *ring drin: ein goldner ring mit brillanten*
 a ring in i:t a golden ring with diamonds

O announces a story about something she read in the paper (l. 1). Before launching into the story, she performs a forward looking repair in that she double-checks whether M is familiar with a reference: the market booths, specifically the butcher’s booth. By uttering an acknowledgement token in l. 4, M claims knowledge of the referent. Despite this, O provides additional descriptions of the referent (l. 5). O produces a third turn repair on the term *markt* /‘market’ which could potentially be heard as a street name or alternatively as an event. It is in response to this self-repair that M produces the multiple. It is produced as one unit of talk in which one element is repeated by the same speaker in immediate succession and under one intonation contour, which is falling (see Figure 2).

Figure 2. A *^jaja* token from the excerpt above.



M conveys that the additional information is unnecessary and that the action of providing additional information can properly be stopped. In fact, M himself provides evidence that he has recognized the referent by providing a descriptor (*untn am dom*/down at the cathedral). Note that this additional description is not try-marked but is instead uttered with falling intonation. O launches into her story, interrupts herself, confirms the additional place name and continues. These findings are in line with Stivers’ [18] research on multiples in English. In Stivers’ data, a unit of talk is repeated in this fashion anywhere from two to seven times with no observed difference in terms of production of and response to a double or a multiple. Stivers argues that multiple sayings do not respond only to the immediately prior utterance, but also respond to the larger action the prior utterance was part of: ‘Multiple sayings function to display that the speaker finds the prior speaker’s course of action problematic, typically its perseveration, and proposes that the course of action be halted’ [18: 288]. In fact, in German, just like in English, speakers produce multiples of this kind for a variety of different tokens.

In terms of its placement, the second type double, i.e., *ja^ja*, is always positioned in interactional environments in which the speakers’ common world view is fractured. The basic sequence unfolds as follows: A produces an utterance and B responds to it. B’s response displays her misalignment with the previous turn. This misalignment can take the form of a challenge, a repair, a sequential misinterpretation, or a pursuit of a previously-given response. It is in response to this

misalignment that Speaker A produces a turn containing a turn-initial *ja^ja* (see Figure 3).

In the following excerpt, T is talking to his sister C. Just prior to this transcript C announced that their father and his new wife will visit C. *Die Krämer* / ‘Mrs. Krämer’, mentioned in l. 5, is a friend of the family.

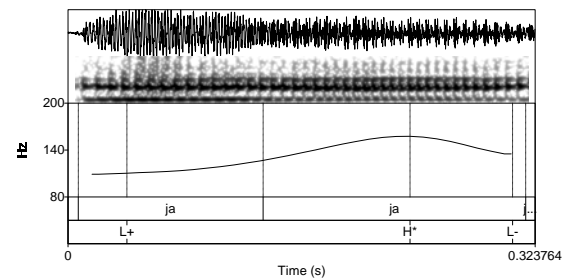
- 1 T: *kuck einer an. von denen hab ich auch scho seit*
look at that. i haven’t heard from them
- 2 T: *e:wigkeiten nix mehr g’hört.*
in a:ges
- 3 C: *echt?* ((unclear since she yawns while talking))
really
- 4 T: *häh*
huh
- 5 C: ((yawn)) *als- hier die krämer hat mir n’en mail*
when- here mrs. krämer sent me an
- 6 C: *geschickt und sie hat jemeint du hast dich bei ihnen*
email and said that you contacted them.
- 7 C: *jemeldet. das war erst vor ↑kurzem,*
that was pretty ↑recently,
- 8 (0.5)
- 9 T: *ich hätte mich bei ihnen gemeldet?*
i contacted them?
- 10 C: *ja.*
yes
- 11 T: => *jaja ich hat ihnen schon en mail geschickt aber von*
yesyes i sent them an email but nothing
- 12 T: *ihnen kommt ja nix und vom [vater hab i ja noch*
is coming from them und from [father i have not
- 13 C: *[echt?*
[really?
- 14 T: *gar nix gehört .hhh [des sin*
heard anything at all .hhh [those are
- 15 C: *[der va:ter is unglaublich*
[father is unbelievable

C’s announcement, that their father and his new wife will be visiting her, is greeted by T with surprise. He complains that he has not heard from his father’s family in a long time. C responds with a marker of disbelief (l. 3). Possibly due to her yawn, which makes this marker hard to understand, T initiates a non-specific repair (l. 4). In response, C accounts for her marker of disbelief by quoting a friend of the family who had informed her that T had been in touch with his fathers family (l. 5-7). This third party quotation is produced as a challenge to T’s prior utterance of not having heard from his father. After a pause of 0.5 seconds, T initiates repair. This repair is done in form of an understanding check; by stressing *ich* / ‘i’, T marks the agent of the verb (i.e., who did the contacting) as the trouble source (l. 9). C confirms this understanding with a simple confirmation marker *ja* / ‘yes.’ Upon this confirmation, T produces a turn-initial *ja^ja*. With his *jaja*-prefaced turn, T confirms C’s understanding that he contacted his father, and points out that his father has not gotten back in touch with him.

T’s turn indicates that he has realized that C took the original utterance ‘i haven’t heard from them’ to mean, ‘i don’t have any news from them’, whereas he meant it as ‘not having been contacted by the father or the father’s family.’ Rather than performing a third position repair [14], [6], which halts the ongoing action, T produces an embedded correction [11] which allows him to propell the ongoing action. This *ja^ja* token does not simply function as an acknowledgment token; instead, the *jaja*-prefaced turn directly deals with C’s

challenge by explaining that it was unwarranted since it was based on a misunderstood fact.

Figure 3. A *ja^ja* token from the excerpt above.



In contrast to the previous segment, in which a *^jaja* was used to indicate the problematicity of a previous action and to halt this action, in the present segment we see a *ja^ja* being used to indicate the problematicity of a previous action and then to deal with it. That is, while a *^jaja* is sequence-closing, a *ja^ja* is not. Moreover, while a *^jaja* can occur in a variety of sequences and contexts, a *ja^ja* only seems to occur when speakers are misaligned, i.e. when their intersubjectivity is fractured. In *^jaja* segments, the coparticipant assumes too little knowledge on the part of the speaker who utters *^jaja*; in *ja^ja* segments, the *ja^ja*-speaker has more information than the coparticipant. Lastly, while a *^jaja* is one particular instance of the practice of multiple sayings which can also be used for a variety of other phrases and tokens, the production of *ja^ja* is its own practice, i.e., it was not observed with other tokens or phrases.

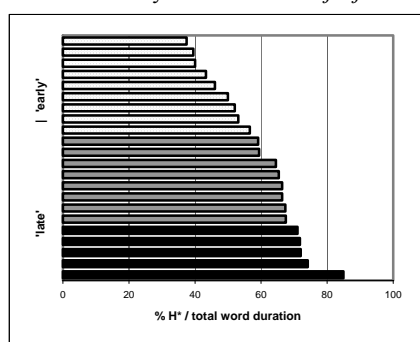
5. ‘Peak accent’ contours in contrast

In light of the CA findings, variable peak alignment observed in the *ja^ja* group could be accounted for as follows. Using GToBI [10], *^jaja* contours, with an H tone in the first syllable and a L tone at the edge of the second, correspond to H* L-% ‘peak accent’ nuclear contours. Such contours have been attested in various unmarked (e.g. non-focus) contexts, e.g., declaratives in GToBI and other models of German intonation (e.g., Féry’s ‘simple falling tone’ in [7: 82]). This interpretation is in line with the wide variety of sequential occurrences of *^jaja* pointed out in the CA account. The *ja^ja* contours, on the other hand, uttered either with a clear L tonal target on the first syllable or a sharp rise or ‘jump’ to an otherwise highly variably anchored H tone on the second syllable, closely match what Grice et al. [10: 65] describe as a ‘rise from low up to peak accent’, L+H* L-%, nuclear contour. The peak is the starred tone (H*) of the bitonal pitch accent, as it is associated with the (second) syllable perceived as high, and it is, indeed, often later in the accented syllable.

As shown in Figure 4, the alignment of H* varies virtually continuously within and across all 22 renditions of *ja^ja* by four speakers from several dialect areas of German. Despite these major sources of inter- and intra-speaker variation, the placement of H* is not random. Six out of nine occurrences of *ja^ja* showing the earliest H* alignment are immediately followed by some segmental material within the phrase, and thus could reflect tonal retraction. Similarly, in all but one out of five cases, the latest H* alignment patterns co-occur with a following silent pause. In four of the early alignment cases, H* is realized at or near the offset of the first syllable (see Figure 1). While one case could be due to some—

albeit very minimal—overlap with a female speaker, in the other cases, speech rate and dialect could have played a role. Two contours were uttered by a female speaker who speaks faster than any of the other speakers. Since rate has not been factored in the measurements of H* alignment, inter-speaker variations (fast vs. slow speakers) might appear as structural (phonological) rather than individual differences. In the fourth contour, pronounced by a Bavarian speaker, H is in the first syllable, preceded by a L target that is longer than in other speakers' renditions of *ja^ja*. The slight fall and 'midish' final tone in the second syllable are also atypical compared to other speakers' *ja^ja* tokens. Although further analyses will be necessary, this could point to different phonetic realizations of *ja^ja* in this speaker's dialect, considering that differences in the realization of rising accents have already been documented between Northern and Southern speakers of German [1].

Figure 4. Alignment of H* with respect to the total duration of the word in twenty-two L+H* L-% *ja^ja* tokens.



Perhaps even stronger support for the analysis of most *ja^ja* tokens as L+H* L-% intonation contours comes from their interactional meaning. As argued above, with the production of such a token, its speaker indicates that the prior action of the coparticipant is unwarranted, in that it is based on wrong or self-evident information. Rather than initiating a repair proper, the speaker uttering this contour acknowledges the coparticipant's turn and simultaneously points out its irrelevance, thereby insisting on his or her epistemic authority. Thus the speaker's 'insistence' motivates the choice of the L+H* L-% delayed 'peak accent' contour, thereby conveying a sharp final fall that has been attributed to exclamations in some inventories of intonation patterns of German [19].

6. Conclusions

The findings of the present study indicate that single sayings and multiple sayings of a given token do not perform identical interactional functions. Moreover, phonetic variants of the same token, in our case two different types of intonation contour, are used by speakers to achieve different interactional goals. This corroborates previous work on response tokens and offers evidence against grouping such items into a single category (e.g., [8], [13]; [17]; [20]).

By showing that the prosodic pattern of an utterance is closely intertwined with the interactional contingencies of a given situation (e.g., [16], [5]), this paper also demonstrates how CA can provide much needed insights into the interactional meaning conveyed by units of intonation phonological analysis. The possible contrast between two 'peak accent' nuclear contours, H* L-% and L+H* L-%, shown in this paper, could have remained hidden in more or less controlled experimental investigations of these tonal

patterns. Just as Local [12] suggested nearly ten years ago, CA and intonation phonology can usefully complement each other in providing new insights and raising new questions on the meaning of intonation in conversation.

6. References

- [1] Atterer, M.; Ladd, R. D., 2003. On the phonetics and phonology of "segmental anchoring" of F0: evidence from German. *Journal of Phonetics* 32(2), 177-197.
- [2] Atkinson, J.M.; Heritage, J., 1984. *Structures of Social Action*. Cambridge: Cambridge University Press.
- [3] Auer, P.; Uhmann, S., 1982. Aspekte der konversationellen Organisation von Bewertungen. *Deutsche Sprache* 10, 1-32.
- [4] Couper-Kuhlen, E.; Ford, C.E., 2004. *Sound Patterns in Interaction*. Amsterdam: John Benjamins.
- [5] Couper-Kuhlen, E.; Selting, M., (eds.), 1996. *Prosody in Conversation*. Cambridge: Cambridge University Press.
- [6] Egbert, M., 2002. *Der Reparatur-Mechanismus in deutschen und interkulturellen Gesprächen*. Verlag für Gesprächsforschung, www.verlag-gesprächsforschung.de
- [7] Féry, C., 1993. *German Intonational Patterns*. Tübingen: Niemeyer.
- [8] Gardner, R., 2002. *When Listeners Talk*. Amsterdam: John Benjamins.
- [9] Golato, A., 2005. *Compliments and Compliment Responses*. Amsterdam: John Benjamins.
- [10] Grice, M.; Baumann, S.; Benz Müller, R., 2005. German intonation in autosegmental-metrical phonology. In *Prosodic Typology*, S.A. Jun (ed.). Oxford: Oxford University Press, 55-83.
- [11] Jefferson, G., 1987. On exposed and embedded correction in conversation. In *Talk and Social Organization*, G. Button, J.R.E. Lee (eds.). Clevedon: Multilingual Matters, 86-100.
- [12] Local, J., 1996. Conversational phonetics: some aspects of news receipts in everyday talk. In E. Couper-Kuhlen, M. Selting (eds.), 177-230.
- [13] Schegloff, E.A., 1981. Discourse as an interactional achievement: Some uses of 'uh huh' and other things that come between sentences. In *Analyzing Discourse*, D. Tannen (ed.). Washington, DC: Georgetown University Press, 71-93.
- [14] Schegloff, E.A., 1992. Repair after next turn: The last structurally provided defense of intersubjectivity in conversation. *American Journal of Sociology* 97(5), 1295-1345.
- [15] Schegloff, E.A., 1998. Reflections on studying prosody in talk-in-interaction. *Language and Speech* 41, 235-263.
- [16] Schegloff, E.A.; Ochs, E.; Thompson, S.A., 1996. Introduction. In *Interaction and Grammar*, E. Ochs, E.A. Schegloff; S.A. Thompson (eds.). Cambridge: Cambridge University Press, 1-51.
- [17] Sorjonen, M.L., 2001. *Responding in Conversation*. Amsterdam: John Benjamins.
- [18] Stivers, T., 2004. "No no no" and other types of multiple sayings in social interaction. *Human Communication Research* 30(2), 260-293.
- [19] Wunderlich, D., 1988. Der Ton macht die Melodie. In *Intonationsforschungen im Rahmen der nicht-linearen Phonologie*, H. Altmann (ed.). Tübingen: Niemeyer, 1-40.
- [20] Zimmerman, D.H., 1993. Acknowledgment tokens and speakership incipency revisited. *ROSLI* 26(2), 179-194.