László Hunyadi: Possible communicative cues to syntactic incompleteness in spoken dialogues
Argumentum 9 (2013), 186-199
Debreceni Egyetemi Kiadó

László Hunyadi

Possible communicative cues to syntactic incompleteness in spoken dialogues*

Abstract

Due to the dynamic nature of dialogues, spoken utterances are often characterized by some degree of fragmentation. This fragmentation is directly reflected by the incompleteness of the syntactic structure of such utterances that represent a considerable challenge for automatic parsing. At the same time, syntactic incompleteness is also a reflection of some cognitive processes underlying the fragmentation of such an interaction. This fragmentation is often accompanied by a set of nonverbal gestures offering thereby the chance to attempt to find certain cues to these cognitive processes by studying the complex of aligning verbal and nonverbal elements. The present paper addresses the issue of the interaction of multimodal markers of conversation with the syntax of spoken utterances to find if certain communicative functions may be cues to the incompleteness of spoken syntax.
Keywords: syntax, spoken utterances, multimodality, communicative functions

1 Introduction

Syntactic incompleteness is a major reflection of the dynamics of conversation. In spoken interaction the verbal performance of the participants is often influenced by several non-linguistic factors, essentially deriving from the mere fact that linguistic planning and production is a function of the cooperative nature of such interactions. Whereas in composing a written text one is generally affected only by one's own cognitive state ranging from thought to verbal form, the cooperative nature of interactions has the effect that the cognitive states of the respective participants also interact. It requires each participant to follow the cognitive state of the other(s) online, incorporating certain elements in their own state and modifying theirs in turn. This is a dynamically evolving process in which the planned and the foreseen go hand in hand with what is emerging. As a result, an already established topic can be modified or can even take a U-turn, the production of an already constructed verbal expression can be interrupted by a sudden turn take of the counterpart, or nonverbal signs can signal a moment of breaking in. Consequently, the individual verbal performance of the participants becomes part of a larger complex: what is ultimately "said" in such an interaction is not a quantitatively incremental sequence of independent individual utterances by the

* This research was partly supported by the following projects: TÁMOP-4.2.2.C-11/1/KONV-2012-0002 and TÁMOP-4.2.2/B-10/1-2010-0024.

alternating partners but a qualitatively new entity emerging from them. It is through this cooperative effort of the partners that the "meaning" of this interaction emerges. What matters here more is the resulting whole rather than its parts: the interaction should be considered well-formed if fulfilling the requirements of having a proper communicative structure, i.e. minimally including a given topic to base the event on and turn management to perform it. On the other hand, the well-formedness of its constituting parts in the form of individual linguistic utterances is subject to the dynamism of the given communicative event and is to be judged as its function. Accordingly, a certain sequence of words may be perceived as rejecting full syntactic analysis (and may therefore be doomed to be ungrammatical), but at the same time such syntactically incomplete expressions as part of a communicative event may be conditioned by the underlying communicative structure and may ultimately be well-formed.

Taking this view of the distinction between syntactic grammaticality and communicative well-formedness in approaching obviously incomplete and fragmented syntactic constructions we can obtain a useful tool for handling them in a systematic way. We need to assume that these constructions are not arbitrary but are the result of some cognitive and communicative processes that leave their trace – at least partly – on the surface in the form of nonverbal markers. Although we leave the question open whether a full enumeration and description of rules belonging to such processes is at all feasible, the study of these communicative markers in relation to verbal expressions can bring us closer to the understanding of the multimodal complex of human interactions.

The HuComTech multimodal corpus[1] of spoken Hungarian of about 60 hours of recordings has the aim to identify those components of human-human communication and their structural relations that are essential building blocks of a communicative event and that can also be implemented in a technological application of human-computer interaction to improve human-machine interfaces. In this connection special attention is paid to the study of the interplay of verbal expressions and nonverbal markers regarding their social and pragmatic functions. Due to the significant share of syntactically incomplete sequences of words in the corpus we implemented a novel scheme of syntactic annotation for spoken language with the aim to capture some of the underlying cognitive and pragmatic processes as reflected in their alignment with nonverbal gestures. Below, we will present some of our preliminary results.

2 A generative model[2] of multimodal communication: primitives vs. markers

In order to identify those components of a communicative event that can serve as relevant cues in the study and interpretation of structural fragmentation of speech, our work is based on a theoretical model of multimodal communication (cf. Hunyadi 2011). According to this theory, the formal description of a communicative event has three main aspects: it needs to be multimodal, modular and generative. Multimodality means that an event, even if mostly based on verbal expressions, is a complex of verbal and nonverbal components. It often happens that a verbal expression is not used in its primary or default sense but it is extended, specified or overridden by some nonverbal form, a piece of prosody or gesture. Modularity[3] involves the expression of certain functions or meanings at various levels: the general frame of all possible

[1] http://hucomtech.unideb.hu/hucomtech.html

[2] The generative concept of this model of communication was inspired by generative theories of language, especially that of Chomsky (1981).

[3] On modularity in pragmatics cf. Kasher (1994).

events is expressed at a purely formal level (in the so-called Basic structure), a finite set of non-contextualized communicative functions at the level of Functional extension, whereas context dependency of particular events at the level of Pragmatic extension. As for generativity, the model operates with the concept of primitives, those minimal ingredients within the particular levels that serve as the basis for generating all and only those events which are considered by our intuition as well-formed within the given context. In turn, these primitives are mapped onto modality specific markers available for surface observation.

3 Unimodal communicative functions

Following this theory, our task then is to identify those surface multimodal markers that can serve as offprints, cues of some structural relations between syntactic incompleteness and observed aligning communicative, pragmatic functions. In what follows, we will restrict our research to the question whether certain communicative functions observed through unimodal, visual markers[4] alone (i.e. without the verbal content conveyed by the audio) show relation to the syntactic structure of spoken utterances. What we wish to learn about is not the way gestures organize the general flow of a conversation alone (cf. Schegloff 1984), but how gestures may serve as cues to some cognitive processes that result in the fragmentation of the syntactic structure of spoken utterances.

We will hypothesize the following:

- the syntactic structure of spoken utterances (completeness, incompleteness and its subtypes) is a reflection of certain cognitive processes;
- these cognitive processes may leave some observable nonverbal cues on the surface;
- the timing between the visual cue and the corresponding speech event (the start of a given clause) may be related to one or more of such cognitive processes.

We will look at a selection of communicative functions as observed through unimodal visual markers and see if any one of them shows a relation to syntactic incompleteness. These functions relate to turn management observable unimodally and include the following: intending to start speaking (short: intending), start speaking successfully (short: success), breaking in (short: break) and end speaking (short: end).[5]

As the names of the functions suggest, we make a distinction between the act of intending to start speaking and the actual performance of starting speaking based on the observation that, often and depending on the dynamics of conversation, the planning of the basic speech act of starting speaking takes some visually observable time and can be differentiated from the actual start of speaking. It is especially important to make such a distinction due to the fact that a dialogue conversation is jointly managed by both interlocutors, and there is often a trading between them for the next turn. In this sense it is understandable that when one of the partners intends to speak, it does not necessarily result in actual speaking, too. We will, however, assume that it is in such trading situations that a special cognitive load is present for the interlocutors that may eventually show up in the corresponding linguistic performance.[6]

In addition, we will also study the possible effect of a sequence of the above communicative functions on syntactic performance: based on the observation that a communicative

[4] Cf. Ekman & Friesen (1969).

[5] For a complete description of unimodal functions see. A. Abuczki (2011).

[6] Hesitation and repair are among those events that have an effect on the fluency of speech; cf. Fromkin (1980), Fox Tree (1995), Gósy (2003). To our knowledge the present paper is the first attempt to study the effect of cognitive events on the syntactic formation of speech utterances in Hungarian in general.

event is virtually always part of a chain of such events we will assume that it develops recursively to yield qualitatively more than just a sum of all previous events. In this sense we expect that certain communicative functions will have a greater effect on syntactic performance in speech if performed in conjunction with others.

4 The annotation of spoken syntax

The still ongoing annotation of our speech material[7] has the following two requirements: a. it should be underspecified with relation to any theory and b. its main focus should be to capture the hierarchical relations between clauses. In addition, we go beyond the sentence level by identifying certain relations (especially elements missing from the surface) through interpretation. Accordingly, whereas we do not annotate formal details of noun and verb phrases, we annotate cases when an argument can be retrieved from a context beyond the given sentence. Accordingly, the basic unit of our annotation is the clause defined as the maximal sequence of grammatically joined words adjacent on the surface. The sentence, in turn, is understood as the maximal sequence of grammatically joined clauses adjacent on the surface. The hierarchy between clauses is annotated in such a way that both the absolute structural relation and their relative surface position is denoted.

The following table describes the taxonomy applied to the spoken syntax of Hungarian in the HuComTech corpus:

Taxonomy of syntactic annotation
Pattern: n.n.n.n.n.n.n. (where n is a one- or two digit number)

Digit #	Definition	Remarks
1	clause ID	the place of the given clause within the sequence of clauses within the sentence ("1" = the first clause of the sentence)
2	subordinating	shows which clause is subordinated to the given clause ("2" = the clause with "2" as its first digit within the given sentence is subordinated to it)
3	coordination	shows which clause is coordinated with it ("3" = it is coordinated with the clause with "3" as its first digit; "3,4" = multiple coordination with clauses "3" and "4")
4	subordinated to	shows which clause the given clause is subordinated to ("1" = it is subordinated to the clause with "1" as its first digit)
5	embedding/ inserting	the clause contains an embedded/inserted clause ("3" = the clause embedded/inserted in the given clause has the ID number "3")
6	embedded/ inserted	the given clause is related to another clause without a grammatical marker ("3a,3b" = the given clause is embedded/inserted in clause "3" where "a" and "b" refer to the two halves of clause "3")
7	categories missing from the surface	implicit categories; specified below

[7] Annotation is being carried out by Hermina Kiss who we wish to thank for her thoughtful work.

	Values	Definition	Remarks
	1	Nothing is missing	The clause is a complete clause.
	2	Head clause is missing	No clause can be identified as its head clause.
	3	Missing clause of coordination before or after the clause	The presence of a conjunction in the clause suggests coordination but no additional clause can be identified as its coordinated clause.
	4	Missing reference word	? Missing relative pronoun?
	5	Missing conjunction	The interpretation of two adjacent clauses suggests that they are related either by coordination or subordination, but the appropriate conjunction is missing from the surface. (Perceived coherence is unmarked.) (‘It started raining. I went home.’)
	6	Missing grammatical subject	The interpretation of the clause points to a grammatical subject which is missing from the surface.
	7	Missing logical subject	The interpretation of the clause points to a logical subject which is missing from the surface.
	8	Missing predicate	The presence of a free adverbial with no further structural relations suggests a missing verb and its arguments.
	9	Missing object	A clause containing a verb whose mandatory object complement is missing from the surface.
	10	Missing adverbial complement	A clause containing a verb whose mandatory adverbial complement is missing from the surface.
	11	Missing attribute	A clause with a nominal phrase missing its mandatory attribute from the surface.
	12	Missing verb	One or more nominal phrases are interpreted as constituting a clause as arguments of a missing verb.
	13	Unfinished clause	In general, the case of corrections or aborted start: more than one obligatory structural unit (specified as 2 through 12 above) is missing.
	14	Complete, but no overt NP-VP structure	Sentential words: Szia! ‘Hi!’ Affirmative answer: Igen. ‘Yes.’ Jó. ‘Good.’ Negative answer: Nem. ‘No.’ Egyáltalán nem. ‘Not at all.’ Uncertain answer: Talán. ‘Perhaps.’ Single word question: Hol? ‘Where?’ Clause starts with the conjunction mint ‘such as’ Politeness marker Conjunction/reference word: Sőt! ‘Moreover!’, És?, ‘And?’, Mert! ‘(Just) Because!’
	15	Grammatical relationship inherently unmarked	insertion

László Hunyadi: Possible communicative cues to syntactic incompleteness in spoken dialogues
Argumentum 9 (2013), 186-199
Debreceni Egyetemi Kiadó

From among the 15 types of syntactic completeness/incompleteness under 7 above we will choose three types with marked differences: type 1 in which none of the obligatory elements of a clause is missing ('nothing missing'), type 13 in which no NP–VP binary structure (the minimal prerequisite of a clause) is present but where these main constituents can be retrieved from a preceding clause (usually the immediately preceding one), and type 14, the case of sentence words where the clause is complete but has no NP-VP structure overt on the surface.

5 Communicative functions vs. syntactic incompleteness: the effect of markers of individual turn functions

Our queries are based on 222 files of formal and informal dialogues conducted with 111 subjects. The subjects were mainly university students aged between 19-28 from the region of Debrecen.[8] The dialogues have been annotated for video, audio, emotions, prosody, syntax and unimodal and multimodal communicative functions. We have queried the 18732 annotations of unimodal functions as well as the annotations of 132853 clauses and 9435 sentences annotated so far for syntax (about one third of the total corpus).

Assuming that there may be a relation between the time of observing a visual marker of a communicative function and the beginning of uttering an aligning clause, we measured their time difference. (The form X > Y shows surface precedence: X precedes Y.) Using the method of principal component analysis (PCA) we wished to find an account for the variation between sets of data in terms of independent factors. The communicative functions included the turn management attributes intending to start speaking, breaking in, start speaking successfully and end speaking; the syntactic types included included the attributes 1 (nothing missing), 13 (more than one argument missing) and 14 (sentence word).
Compare the following data:

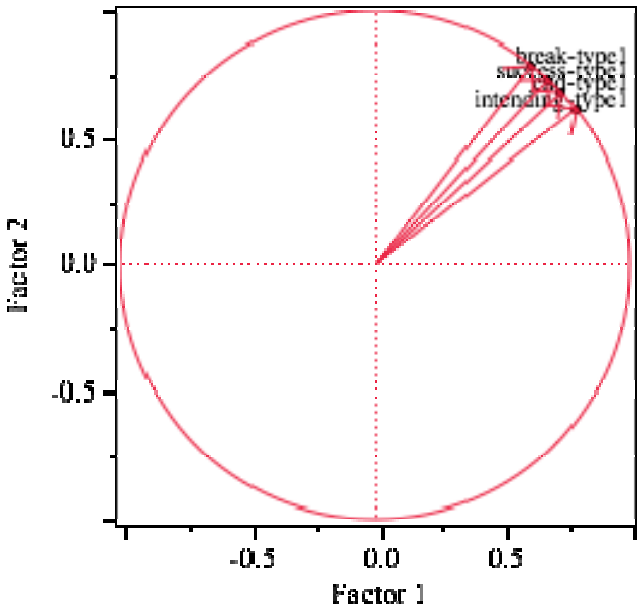

Figure 1: clause type 1 > turn function

[8] For a full list of descriptive data on the corpus see Pápay et al. (2011).

László Hunyadi: Possible communicative cues to syntactic incompleteness in spoken dialogues
Argumentum 9 (2013), 186-199
Debreceni Egyetemi Kiadó

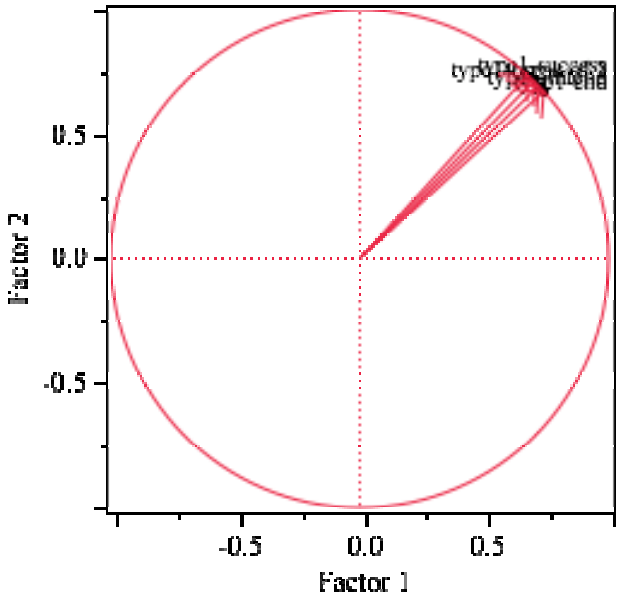

Figure 2: turn function > clause type 1

The graphs above show that the turn functions form a single group in both precedence orders, accordingly, in these cases the completeness of the clauses is not dependent on the turn function.

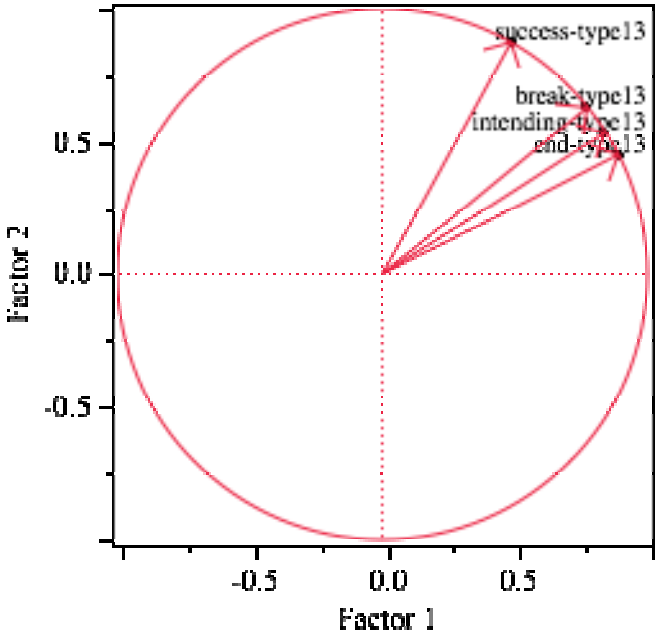

Figure 3: clause type 13 > turn function

László Hunyadi: Possible communicative cues to syntactic incompleteness in spoken dialogues
Argumentum 9 (2013), 186-199
Debreceni Egyetemi Kiadó

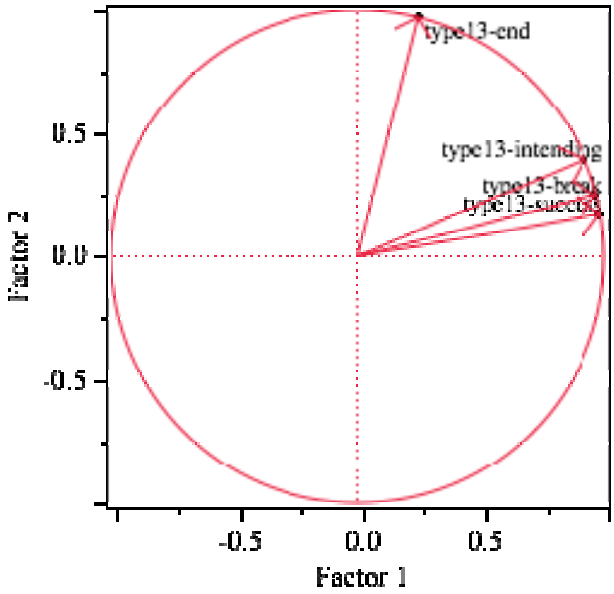

Figure 4: turn function > clause type 13

As the graphs show, in the case of type 13 the functions group into two factors: it is success that stands out in the order type 13 > turn and end in the opposite order turn > type 13.

That the factor pointing to a single variable is different in the two different (opposite) types (sequences) indicates that turn management is directed to different clauses: when turn management precedes the clause it is directed to that clause, if the clause precedes the turn management marker, it is directed to the following clause (that can be either the next one by the speaker or by its counterpart).

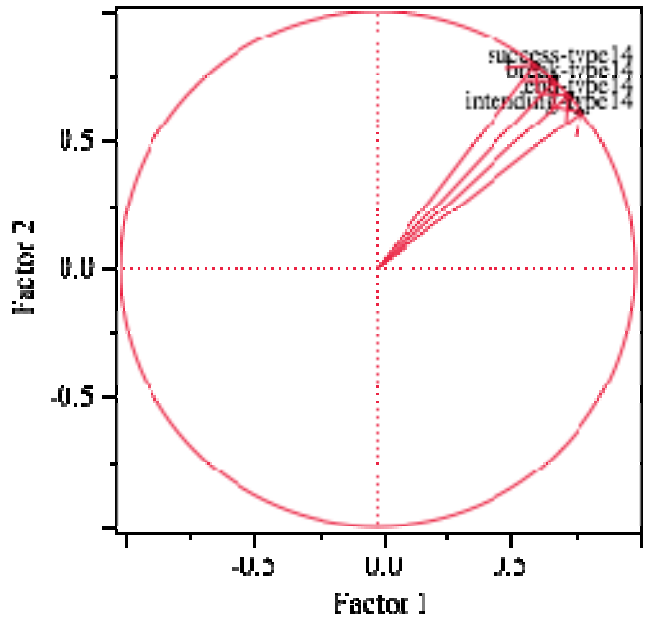

Figure 5: clause type 14 > turn function

László Hunyadi: Possible communicative cues to syntactic incompleteness in spoken dialogues
Argumentum 9 (2013), 186-199
Debreceni Egyetemi Kiadó

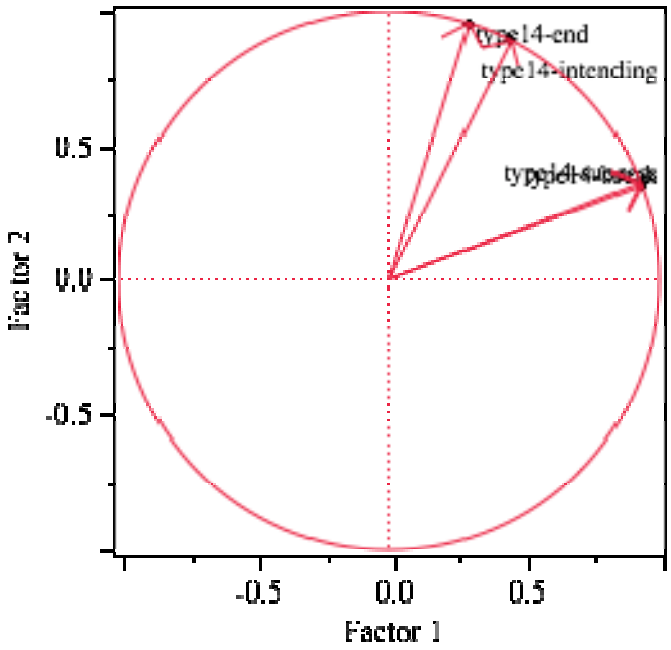

Figure 6: turn function > clause type 14

Depending on the precedence of the visual marker and the clause, the above two graphs show a difference again: whereas the variation of the given functions cannot be related to more than one factor in case the clause precedes the visual marker, the same functions in the opposite configuration (the turn function precedes the clause) are grouped according to two different factors: ending and intending on the one hand, and break and succeed on the other.

Although we must be aware that a. we do not expect a given cognitive function to be represented unambiguously by any linguistic form, let it be syntactic or else, and b. the observed grouping differences are not quite strong to point to one or another factor exclusively, we observe the following tendencies:

We get slightly higher values for possible factors in the configurations where the turn management marker precedes the given clause as compared to the opposite sequence, where the clause precedes the marker. It is justifiable by the fact that a visual marker usually precedes speech production (cf. Allwood and Ahlgrén 2009); accordingly, we can expect a stronger relation between a marker with the clause that it precedes.

In the case of clause type 1 (clauses containing the appropriate surface representation of the basic NP-VP pair and all the obligatory arguments) the data show that, even though all the four kinds of functions are observed prior to the utterance of the clause, there is no difference between them in the time elapsed between the onset of the marker and the start of the clause. Assuming that the elapsed time might reflect the cognitive effort needed for the production of a clause in speech, the fact that all four functions were observed with similar timing indicates no cognitive difference in the production of the clauses. Since clause type 1 is the case of complete syntactic structure, the similarity in the timing of the data suggests that in these cases the different turn functions did not put a special load on language production.

The case of type 13 (clauses with more than one obligatory argument missing on the surface), however, shows the effect of this cognitive effort: in the order when the clause precedes the observed visual marker, factor 2 (with an effect on a single turn function) has an outstanding effect on the function “start speaking successfully”: it is the case when, due to hesitations, breaking in and other rapid interactions with the counterpart resulting in an

incomplete clause (clause type 13) the speaker makes a special cognitive effort to repair his words or continue expressing his thought. This is the effect that can be observed in this sequence.

In the other precedence form regarding clause type 13, i.e. when the visual marker of turn management precedes the given clause, we again have an outstanding function as an effect of factor 2, this time that of "end speaking". Here, there is again some cognitive event taking place as a result of which the speaker ends speaking without completing the clause itself (NB.!: ending speaking is not the same as ending a clause; for several reasons, such as hesitation, breaking in by the counterpart, changing one's train of thoughts and many others, one may not finish a given syntactic structure.) This is what is reflected by the role of factor 2 as effecting "end speaking". The rest of the turn functions group together as an effect of factor 1 with a generic effect with the meaning "speaking" only.

The case of clause type 14 (mostly containing a single word –"sentence words") behaves differently depending on the sequence of clause and turn marker. When the clause precedes the marker, the marker is not specific to the clause: it is justifiable since these one-word clauses ("Yes", "Indeed", "No", "Why" etc.) are usually reflections of the immediately preceding clause; accordingly, the subsequent visual marker is instead directed to the next clause). On the other hand, when the visual marker precedes the clause (i.e. when we expect the marker to be directed to the subsequent clause in questions), factor 1 groups the functions "break-in" and "start speaking successfully" and factor 2 groups the functions "intending to start speaking" and "end speaking" together . We suggest that, in this case, factor 1 is related to a rapid, "marked" cognitive action, whereas factor 2 is related to a more relaxed, default or "unmarked" cognitive action.

6 Communicative functions vs. syntactic incompleteness: the effect of a sequence of markers of individual turn functions

We looked at three-annotation sequences, i.e. sequences where three turn function annotations follow one after the other, and we wanted to find out what syntactic types they align with. The turn functions in question were the same as above, and the syntactic types were those of type 1 (complete clause), type 13 (incomplete clause with more than one argument missing) and type 14 (sentence words). We looked at two configurations: a given, predefined function followed by two functions of any kind (any 1 and any 2), and, in the opposite direction, two functions of any kind (any 1 and any 2) followed by a given, predefined function. We expected that these two different directions may shed light on the difference between turn functions as related to one or another syntactic type.

In addition, we defined two kinds of a given sequence pattern: one in which it was expected that each function time aligns with a clause, and one in which we were just looking at what types of clauses follow (precede, depending on the precedence order) the clause aligning with the given, predefined function, without requiring a strict alignment.

The first finding was a support for the possible expectation that turn functions, i.e. visual markers of intending, starting, ending speaking or breaking in do not necessarily coincide with any speech utterance (they actually coincide in only 7 to 22% of the cases). A given function followed by any other function: intending 22%, breaking in 17.45%, start speaking successfully 5.63%, end 0%; a given function preceded by two "any" functions: break 13.6%, intending 12.05, success 7.2%, end 0. These observations are also indications of the independent role of

visual markers in following a verbal interaction[9]: as the above data show, around 80% or more of conversation in our corpus is accompanied by visual markers of turn management alone, markers that bring forward the interaction preceding, following or beyond its verbal part.

Looking at cases when each of the turn functions in the sequence were aligned with a clause, we found the following:

intending > any > any (23)
most frequent pattern:
intending (23) > break (23) > intending (22)
types of clauses aligning:
clause 1: 1 (0) 13 (1), 14 (7)
clause 2: 1 (1) 13 (0), 14 (3)
clause 3: 1 (1) 13 (1), 14 (4)

The large share of clause type 14 (sentence word) may be an indication of the dynamics of the conversation where break is followed by a short utterance expressing the intension to speak again.

break > any > any (26)
exclusive pattern:
break (26) > intending (26) > break (26)
types of clauses aligning:
clause 1: 1 (3) 13 (0), 14 (15)
clause 2: 1 (1) 13 (1), 14 (9)
clause 3: 1 (3) 13 (1), 14 (2)

This pattern is the opposite of the one above, but the dynamics is similar: clause type 1 (complete clause) is followed by type 13 (more than one argument missing), then followed by the sentence word type 14. This sequence, especially the dominant presence of clause type 14 across the pattern shares similar dynamics with the case of intending above.

success > any > any (9)
most frequent pattern:
success (9) > break (8) > success (7)
types of clauses aligning:
clause 1: 1 (0) 13 (0), 14 (7)
clause 2: 1 (1) 13 (0), 14 (3)
clause 3: 1 (0) 13 (1), 14 (4)

The fact that the function success is followed by break to be followed by success again suggests a strong dynamics of the interaction; hence the predominance of clause type 14.

any > any > intend (23)
most frequent pattern:
intending (21) > breaking (23) > intending (23)

[9] For early observations regarding the role of gestures in organizing the flow of interactions cf. Duncan (1972), Wiemann and Knapp (1975), Duncan and Fiske (1977).

László Hunyadi: Possible communicative cues to syntactic incompleteness in spoken dialogues
Argumentum 9 (2013), 186-199
Debreceni Egyetemi Kiadó

types of clauses aligning:
clause 1: 1 (0) 13 (0), 14 (7)
clause 2: 1 (0) 13 (0), 14 (0)
clause 3: 1 (0) 13 (0), 14 (0)

The dynamics of the interaction is reflected again by the large share of clause type 14 (sentence word) and the lack of complete clauses (type 1) across the pattern.

any > any > break (26)
exclusive pattern:
break (26) > intending (26) > break (26)
NB.! It reflects the increased dynamics of the conversation.
types of clauses aligning:
clause 1: 1 (3) 13 (0), 14 (15)
clause 2: 1 (1) 13 (1), 14 (9)
clause 3: 1 (3) 13 (1), 14 (2)

The high number of type 14 clauses may be a fair indication of the dynamics of the given stretch of conversation involving a rapid alternation of break and intending.

any > any > success (7)
most frequent pattern:
success (6) > break (6) > success (7)
types of clauses aligning:
clause 1: 1 (3) 13 (1), 14 (2)
clause 2: 1 (1) 13 (0), 14 (1)
clause 3: 1 (0) 13 (0), 14 (0)

The lack of type 13 and type 14 clauses with a higher share of type 1 may correspond to the aligning success function implying completeness.

We have to emphasize here, however, that these analyses may only be modest approximations of the actual processes behind the visual observation of the annotator-observer and the speech production of the speaker alike. The annotator did not hear the verbal content and relied on stereotypes associated with the visual gestures. His/her interpretation was, however, assisted by the knowledge of the incremental nature of sequences of gesturing in communication that significantly narrows down possible choices. As for the identified clause types associated with the predominantly dynamic flow of interactions, the small number of type 1 (complete) clauses might meet our expectations, however the equally small number of incomplete clauses (type 13) may be surprising. However, in view of the high number of type 14 clauses it may be suggested that this very kind of interaction with continuous speaking (without the shortest pause) between the clauses does have this very particular property: speak short, break in rapidly and start speaking – in any of the possible orders. We would probably expect a different distribution among syntactic types (including type 13) in a conversation with slower, more moderate interaction.

7 Summary

The aim of the paper was to address the question of possible relations between communicative function and linguistic form with a further aim to get closer to the understanding of certain cognitive processes behind linguistic production in speech. We chose to examine the unimodal functions related to turn management intending to start speaking, breaking in, start speaking successfully and end speaking. The clause types concerned were related to completeness (type 1) and incompleteness (type 13) and also included the special type 14 of sentence words. We found that type 13 is related to the communicative function of end speaking: since type 13 represents a clause with more than one argument missing, consequently, such a clause has a quite fragmented syntactic structure, therefore the indication of ending speaking can easily be associated with certain prominent visual gestures that differentiate this function from the rest of the functions studied here. We also found that the functions intending to speak and ending speaking preceding a clause of type 14 (the clause usually consisting of a single word) group together as an effect of a separate factor: it again shows the possibility that, due to the special nature of clause type 14 (short in form, usually a brief reflection to the communication of the interlocutor) has an accompanying set of visual markers that are distinct from breaking in and start speaking successfully. We need to study further, however, what the content of clauses of type 14 is in order to get stronger generalizations.

As for the cognitive processes behind all this syntactic variation and the interpretations of the accompanying visual gestures, we may conclude that we have got somewhat closer to these relations by at least witnessing that the given turn functions are not arbitrarily distributed across the aligning clauses of different syntactic types. Strongly believing in the multimodal nature of both verbal and nonverbal communication we can expect that these crossmodal relations do exist and are essential components of communication and can even account for important properties of certain formal variation. However, due to this very nature of communication and language production we need to conclude that in order to have an even deeper understanding of these properties we need to extend the range of our research objects, go beyond the present means and methodologies. That is how, by steady small steps, we might get closer to the understanding of our cognition that actually controls our language, speech and communication.

References

Abuczki, A. (2011): A multimodal analysis of the sequential organization of verbal and nonverbal interaction. *Argumentum* 7 (2011), 261-279.

Allwood, J. & Ahlgrén, E. (2009): Gestures that precede and accompany speech – An analysis of their functions and use in the design of virtual agents in different activities. In: Navarretta, C., Paggio, P., Allwood, J., Alsén, E. & Katagiri, Y. (eds.): *Proceedings of the NODALIDA 2009 workshop Multimodal Communication –from Human Behaviour to Computational Models*. NEALT Proceedings Series, Vol. 6 (2009), 3-8.

Bunt, H. & Black, W. (2000): The ABC of Computational Pragmatics. In: Bunt, H. & Black, W. (eds.): *Abduction, Belief and Context in Dialogue: Studies in Computational Pragmatics*. Amsterdam: John Benjamins, 1-46.

Chomsky, N. (1981): *Lectures on Government and Binding: The Pisa Lectures*. Holland: Foris Publications.

Duncan, S. (1972): Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology* 23.2, 283-292.

Duncan, S. & Fiske, D. (1977): *Face-to-face interaction: Research, methods, and theory.* Hillsdale: Lawrence Erlbaum Associates.

Ekman, P. (1957): A methodological discussion of nonverbal behavior. *Journal of Psychology* 43, 141-149.

Ekman, P. & Friesen, W.V. (1969): The Repertoire of Nonverbal Behavior: Categories, Origins, Usage and Coding. *Semiotica*, 49-98.

Fox Tree, J.E. (1995): The effect of false starts and repetitions on the processing of subsequent words in spontaneous speech. *Journal of Memory and Language* 34, 709-738.

Fromkin, V.A. (ed.)(1980): *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand*. Academic Press. New York.

Gósy, M. (2003): A spontán beszédben előforduló megakadásjelenségek gyakorisága és összefüggései. *Magyar Nyelvőr* 127, 257-277.

Hunyadi, L. (2011): Multimodal human-computer interaction technologies. *Argumentum* 7 (2011), 240-260.

Kasher, A. (1994): Modular Speech Act Theory: Research Program and Results [collection on the foundations of speech act theory] In: Tsohatzidis, S.L.T. (ed.): *Foundations of Speech Act Theory Philosophical and Linguistic Perspectives.* London & New York: Routledge, 312-322.

Kopp,S., Bergmann, K. & Wachsmuth, I. (2008): Multimodal communication from multimodal thinking – Towards an integrated model of speech and gesture production. *International Journal of Semantic Computing* 2.1, 115-136.

Pápay, K., Szeghalmy, Sz. & Szekrényes, I. (2011): HuComTech Multimodal Corpus Annotation. *Argumentum* 7 (2011), 330-347

Schegloff, E.A. (1984): On Some Gestures' Relation to Talk. In: Atkinson J. M. & Heritage J. (eds.): *Structures of Social Action: Studies in Conversational Analysis*. Cambridge: Cambridge University Press, 266-296.

Wiemann, J. & Knapp, M. (1975): Turn-taking in conversations. *Journal of Communication* 25, 75-92.

László Hunyadi
University of Debrecen
Department of General and Applied Linguistics
H-4010 Debrecen
hunyadi@ling.arts.unideb.hu