THE LIMITS OF GRAMMAR:
CLAUSE COMBINING IN FINNISH AND JAPANESE
CONVERSATION[1]

Ritva Laury and Tsuyoshi Ono

Abstract

Our paper concerns the grammar of clause combining in Finnish and Japanese conversation. We consider the patterns of clause combining in our data and focus on the verbal and non-verbal cues which allow participants to determine whether, after the end of a clause-sized unit, the turn will end or continue with another clause-sized unit, resulting in a clause combination. We conclude that morphosyntax alone cannot account for the patterns found in our data, but that the participants orient to, at least, prosodic and non-verbal cues in determining the boundaries of clauses and projecting continuation in the form of another clause. Also important for projection are fixed expressions or 'prefabs'. In addition, semantic and pragmatic factors play a role. In that sense, we explore the question of where the limits of grammar for interaction, understood as the knowledge which speakers share and which forms the basis for the creation and processing of novel utterances, should be drawn, and whether grammar should include, beyond morphosyntax, not only prosodic, pragmatic and semantic features but also bodily behavior.

Keywords: Clause combining; Japanese; Finnish; Morphosyntax; Prosody; Non-verbal behavior; Prefabs.

1. Introduction

Our paper concerns the grammar of clause combining in Finnish and Japanese conversation. As proponents of usage-based, discourse-functional and emergent grammar, we build on the claim that grammar is "the cognitive organization of one's experience with language" (Bybee 2010: 8), or, put slightly differently, "the collective sum of actual speakers' experiences which is seen as the basis of the creation of novel

utterances" (Auer & Pfänder 2011: 4). This conception of grammar has the obvious consequence that grammar emerges, that is, comes into being, during instances of language use (Hopper 1987, 2011). This is necessarily so, of course, if grammar is based on, or is a sum of, experience.[2] However, the notion of emergence does not entail that grammar has no stability. On the contrary, many patterns in what we call grammar appear to be stable and indeed shared by speakers in spite of the fact that everyone's experience with language, the source of grammar as we understand it, is different. The stable patterns range from lexically specific, idiomatic set expressions to relatively abstract patterns. The stability of patterns, in turn, is directly related to, and results from, their frequency of use (see, e.g. Bybee & Hopper 2001; Du Bois 1987). Methodologically, this leads to an approach where we as analysts find patterns in actual language use and present them as hypotheses about knowledge/skills (representation, or grammar) that allow people to produce language in actual interaction. For the linguist, capturing these patterns properly has been shown to require attention to both structure and use, and more recently also to non-verbal dimensions in the case of data from face-to-face conversation (see, e.g. Goodwin 2003; Haddington, Mondada & Nevile 2013).

The clause has long been considered a central unit of language by many linguists of varied theoretical backgrounds, and the combining of clauses is a topic of constant interest in accounts of the structure of human languages. Within interactional linguistics, Thompson and Couper-Kuhlen have suggested that the clause is "the favorite unit of linguists" (2005: 484), and Helasvuo (2001) considers the clause a central resource in interaction. In their classic paper, Sacks, Schegloff and Jefferson claimed that in addition to clauses, sentences, which we take to mean clause combinations, are one of the building blocks of Turn Constructional Units (1974: 721). At the other end of the spectrum, proponents of autonomous syntax have even claimed that recursion, the embedding of similar structures such as clauses within one another, is the defining feature of human language (Hauser, Chomsky, and Fitch 2002; for an opposing argument, see Laury and Ono 2010). It is indeed true that in terms of grammars of individual languages, one would be hard put to find a reference grammar that would not deal with the topic of clause combining, and the ways in which clauses are combined in complex sentences is a classic parameter of linguistic typology (e.g. Keenan and Comrie 1977). However, linguists working on spoken data have noted that while clauses seem to emerge quite clearly as a gestalt in spoken data from many languages, the idea of the sentence may be a linguist's category, reflecting perhaps the written language bias in linguistics (Linell 2005). A number of linguists have questioned the suitability of the concept 'sentence' for the analysis of spoken data (e.g. Miller 1995; Iwasaki & Ono 2001). For example, it has been shown that the determination of sentence boundaries is difficult in spoken data, since speakers regularly add material ('increments'; Ford, Fox & Thompson 2002; Couper-Kuhlen & Ono 2007) to already completed utterances. Such expansions may be fitted to what precedes, or they may even reanalyze the preceding syntax (e.g., Tanaka 2001).

The main finding arising from our data is that clause combinations in Finnish and Japanese conversation are put together by producing and connecting clauses one by one, on-line (Auer 1992, 2005), in interaction with the other participants. We further argue that the final specific shapes of multi-clause combinations are an adventitious outcome resulting from responding to various on-line factors. We thus suggest that the resulting

---

[2] Both of these ideas, of course, stand in contrast to the stance common in autonomous linguistics that language use does not affect language structure (see, e.g., Newmeyer 2003).

multi-clause combinations with specific lexical items do not become permanent structures for the participants; that is, clause combinations do not seem to become set expressions, constructions or 'prefabs' the way that certain clause types do. In fact, we will suggest that what may seem to the analyst *a posteriori* a clause combination, a structured unit, is not, in our view, a structured, pre-planned grammatical unit for the speakers. Rather, speakers of Finnish and Japanese only seem to know how clause-sized units are built and how to combine them one by one, on-line, at the local level, using a set of resources including clausal grammar, clause connecting devices, conventionalized formulas, prosody, and gesture. However, as we will show in what follows, each multi-clause combination may not form a coherent whole, and does not appear pre-planned. Nor does it necessarily conform to descriptions of clause combinations (often called sentences) in reference grammars.

We arrive at these findings, in the spirit of usage-based grammar, by looking for instances of combining clause-sized units in language use. The central evidence for us of a shared grammatical pattern is its production and recognition by participants in conversation. In particular, we are interested in what participants share that allows them to recognize whether and when a turn is about to end after a clause-sized unit, allowing another participant to take a turn, or whether it is going to continue, leading to a clause combination consisting of a new clause combined with the previous clause(s).[3] This perspective on clause combining is consistent with the view of grammar as "a conventionalized set of formal ways of making projection possible" (Auer 2009: 180). We look for how such recognition, the shared knowledge of the grammar of clause combining, is manifested in the data. While we are not able to make strong claims regarding conventionalization of the patterns we present below mainly due to the lack of diachronic data, the patterns are representative of the many cases of clause combinations we have examined (see also, e.g., Laury & Ono 2010, which was based on a 2000-clause random sample corpus in Finnish and Japanese). Thus, although each multi-clause combination in our data is arrived at contingently, emerging in the local context in response to various factors, the set of means used for projecting continuation is conventionalized, to the degree that it is used by speakers and recognized by the other participants, and in that sense, grammatical.

What, then, is grammatical about the clause combinations in our data? In projecting the ends of turns, speakers do rely on morphosyntactic patterns in clausal grammar such as argument structure and verb forms. In fact, it does seem that in our data, as suggested by Thompson and Couper-Kuhlen (2005), the clause is one of the most frequent formats the speakers orient to, and that, as suggested by Helasvuo (2001) "the clause is precisely that unit which permits significant projectability" (485).

What also guides participants in projecting whether a turn will end after a clause-sized unit are set expressions, or constructions that may be entirely or partly lexically specific. As has been shown, especially certain main clause types become conventionalized (formulaic), projecting more talk of a particular type (e.g. Scheibman

---

[3] Clauses and TCUs can, of course, be continued after a completion point in a variety of ways (see, for example, papers in Couper-Kuhlen & Ono 2007). In this paper we are dealing specifically with cases where clauses are combined with other clauses. We consider linkages done with conjunctions, but also ones without conjunctions, for example, linkages done only with prosody. It is of course true, as one of the editors has pointed out to us, that participants combine TCUs, turns and actions, but we are here specifically concerned with the combining of clausal TCUs, while TCUs, of course, can take of a variety of syntactic shapes. For discussion of clause combinations vs. TCU combinations, see Couper-Kuhlen 2012; for a discussion of participant orientation to clauses, see Thompson & Couper-Kuhlen 2005.

2001; Keevallik 2003; Kärkkäinen 2003; Couper-Kuhlen & Thompson 2008; Hopper & Thompson 2008; Pekarek Doehler 2011; Kärkkäinen forthcoming). Beginnings of turns often include phrases that semantically and pragmatically project a certain type of clause, or a certain type of discourse to follow, such as a quote, a story or an account. Central among the devices that seem to guide participants in the way we describe above are also features of talk at the ends of clause-sized units, in particular conjunctions and particles.

In our data, as well, besides clause structure, conventionalized formulas and particle-sized units, prosodic features of talk and bodily behavior, including behavior related to other activities at hand at the time, also appear to play an important role in projection and in shaping clause combinations. In fact, much clause combining in our data appears to be responsive to what the other participants are doing in terms of minimal backchanneling and ratification. The shape that the combinations take, the adding of clauses and finishing up the clause combination, has much to do with what the other participants have done just before or are doing at the point in time that the turn is underway. In that sense as well, clause combinations in our data are emergent, contingent, and co-constructed by the participants.

While many of the combinations as a whole in our data might seem to involve some ungrammaticality if one takes as a point of departure standard descriptions of Finnish and Japanese grammar, they contain features which speakers seem to operate with and thus are shared for the purpose of combining clauses one by one, and in that sense conventionalized, fitting within our understanding of grammar.

Much research in discourse-functional linguistics and in interactional linguistics has been devoted to patterns that can be said to have already emerged or to be emerging in the sense of being in the process of becoming grammaticized, conventionalized or partially fixed. In our clause combining data, there seem to be few combinations that can be said to have already become conventionalized or even partially fixed. The goal of our article is to determine the limits of grammar – what is shared among speakers, and what is not, in clause combining, and which features of behavior, verbal and non-verbal, should be included within the grammar of clause combining.

2. Data

Our Finnish data come from the collection of face-to-face and telephone conversations housed in the Conversation Analysis Archives in the Department of Finnish, Finno-Ugric and Scandinavian Studies at the University of Helsinki. The Japanese examples are selected from a set of conversational recordings and transcripts collected and prepared at UC Santa Barbara and UCLA around 1990, currently housed in the Spoken Discourse Research Studio at the University of Alberta.

Our examples are presented based on the transcription system originally developed by Chafe (1980) and later refined by Du Bois et al. (1993); for explanations of the symbols used, see the appendix. The core of this transcription system is the identification of an auditory unit called the intonation unit, which is defined as a stretch of talk with a coherent pitch contour (Chafe 1987, 1994; Du Bois et al. 1993; for Finnish, see e.g. Aho 2010; for Japanese, see Iwasaki 1993a). This auditory impression is often created (thus perhaps supported) by the presence of a pause and a pitch reset between intonation units, both of which are recognized as cues for the identification of

intonation units in the stream of speech. It has been suggested that intonation units are found in all languages (Chafe 1987, 1994; Du Bois et al. 1993). The division of the examples into intonation units in this paper was reached through consensus following independent checking by multiple transcribers. Each line in our examples represents an intonation unit. For our purpose, features such as gaze and head nods, which are not originally in the Du Bois/Chafe system, are also transcribed.

In what follows, we first discuss and exemplify the structure of Japanese and Finnish clauses and clause combinations, as described in reference grammars, for the benefit of those readers who might not be familiar with one or both of the languages. We then offer a micro-analysis of three typical data segments in our data. We go over our examples in a modular[4] (or compartmentalized) way, focusing on one area of language at a time. Having tried to describe conversational segments, and develop hypotheses about what is going on in them for a number of years, it is becoming increasingly clear that picking which factors to highlight and balancing the discussion of multiple factors requires a theory of its own, which we currently do not have. Though we think that human language actually operates in a non-modular way, and that by discussing it in a modular way we run the risk of distorting what is really happening, we still chose a modular approach in describing and analyzing the examples in this paper. This is simply because it made it possible for us to examine and, equally importantly, discuss all the potentially relevant factors exhaustively.

3. The clause in Finnish and Japanese

3.1. *Finnish*

Finnish clauses are, in their main features, quite similar to those in well-known Indo-European languages such as English, except for the fact that Finnish has an extensive morphological case system. For the definition of the clause, Finnish reference grammars rely on the notion of the finite verb, which is the only necessary component of the clause. Finite Finnish verbs are marked for person, and they may also contain inflectional morphology for tense and mood. In addition to the verb, the clause core may include a subject, an object, or an adverbial complement (Hakulinen et al. 2004: 827-830). Although Finnish has, in principle, free word order, so that no order of clausal components results in ungrammaticality, the neutral clausal order is SV(X), and it is also the dominant order in conversational Finnish data (Helasvuo 2001: 79-81). In spoken Finnish, separate pronouns are used in addition to the person marking in the verb (see Helasvuo, this volume); in contrast, in written Finnish, first and second person referents are normally not coded with separate pronouns.

In Finnish clause combinations, subordinate clauses follow their main or matrix clauses: Relative clauses come after their heads, and complement clauses normally follow their matrix clauses. Adverbial clauses are, as one might expect, more free in terms of occurrence; they may precede or follow their main clauses. Relativizers, complementizers, and adverbial conjunctions occur before the subordinate clause. Differently from languages like Swedish and German, for example, word order does not

---

[4] One of the Pragmatics anonymous reviewers suggests not using the term 'modular' since it is associated with a particular view of language in the field of linguistics. We specifically are not claiming that language is modular, as should be clear from the way we use the term here.

distinguish between main and subordinate clauses in Finnish. And, as will be shown in this paper, clauses that are formally subordinate and marked as such with subordinating conjunctions may appear independently, without any main clauses (Laury 2006; Koivisto, Seppänen & Laury 2011; Laury 2012; for a typological account of 'insubordination', see Evans 2007). The following examples illustrate Finnish clauses and clause combinations. They all come from the Comprehensive Grammar of Finnish (Hakulinen et al. 2004). Example (1a) contains a relative clause, example (1b) a complement clause and (1c) a (preposed) adverbial clause.

(1a)
Valitse   maa,   jo-sta   ole-t   kiinnostunut.
choose.2SG.IMP   country.NOM   REL-ELA   be-2SG   interested
'Choose the country in which you are interested.'

(1b)
Sano-i-n   hei-lle, että   me ole-mme liian pien-i-ä.
say-PST-1SG 3PL-ILL COMP 1PL be-1PL   too small-PL-PRT
'I told them that we are too small.'

(1c)
Kun kolme-kymmen-tä vuot-ta sitten ryhdy-i-n   kerää-mä-än   aineisto-a,
when three-ten-PRT   year-PRT ago start-PST-1SG collect-INF-ILL data-PRT
luul-i-n   tehtävä-ä helpo-ksi.
think-PST-1SG task-PRT   easy-TRA
'When I started to collect data thirty years ago, I thought the task (would be) easy.'

3.2. *Japanese*

Japanese clauses look rather different from Finnish clauses because they involve a final predicate, as illustrated with constructed examples from (Japanese) linguistics textbooks and grammars given below. These are all single clauses and represent the so-called 'sentence' in Japanese, which is characterized by NPs marked with postpositional 'case' and ending with the predicate, which is tensed but not marked for person. Such examples are abundant particularly in introductory discussions of the language. It seems fair to say that such examples are responsible for having created a widely held view of what Japanese is like.[5]

---

[5] These examples contain inconsistencies in phonological and morphological analysis, gloss, and translation compared to other examples cited in this paper because they come from various published sources. We have retained the original glossing in examples (2) and (3) because they are results of competing analyses which have not been resolved. This is partly because many linguists borrow analytic categories from descriptions of languages very different from Japanese. Consider, for example, the glossing of the nominal particles such as *ga* and *o,* which are sometimes glossed with case labels such as 'nominative' and 'accusative', sometimes with syntactic labels such as 'subject' and 'object', while they are left entirely unglossed by some authors. There is also inconsistency as to whether they are thought to be cliticized or affixed to, or independent from, the noun they follow, shown in the way the hyphen is used or not used. In the glossing of our own examples, we have simply used GA and O, since there is no agreement among scholars on the nature and thus glossing of these morphemes.

(2a)
Taroo-ga ringo-o tabeta
Taroo   apple   ate
'Taroo ate an apple.'   (Stewart & Vaillette 2001: 193)

(2b)
watashi ga   sakana o   tabete   iru
I   subject marker fish   object marker ea(ting) am
 'I am eating fish.   (Fromkin & Rodman 1993: 121)

(2c)
Taroo ga   kita
          NOM came
'Taro came.'   (Shibatani 1990:260)

(2d)
Taroo-ga   kinoo   ookii hanbaagaa-o   itutu tabeta
Taro-Nom   yesterday big   hamburger-Acc five ate
'Taro ate five big hamburgers yesterday.'   (Tsujimura 1996:160)

(2e)
fuji-san ga   utsukushii
Mt. Fuji NOM beautiful
'Mt. Fuji is beautiful.'   (Iwasaki 2002:85)

     Though predicate finality is relatively rigid, the order of NPs before the predicate is free, and this flexibility is often said to be due to the presence of postpositional 'case' marking the role of the NPs in the clause. As we will see below, however, NPs are actually often found without postpositional 'case' in conversation, which casts doubt on the suggested connection between the flexibility of constituent order and the presence of postpositional 'case'.

     Perhaps the most significant aspect of Japanese clauses is that NPs are not required elements. So, for instance, the clauses in (2) would all be acceptable without some or all of the associated NPs, a pattern which is very commonly observed in conversation. Thus, no overt reference is required, even though Japanese verbs are not person marked.

     As we will see below, the literature also abounds with constructed examples where two clauses that look similar to the above are combined by a postposed subordinate clause marker to form a clause combination, again commonly called 'sentence', as can be seen in examples 3a-d below.

(3a)
John ga kitara   boku wa kaeru
          come-if I   leave
'If John comes, I will leave.'   (Kuno 1973: 177)

(3b)
John ga Mary o nagut-te,[6] *boku ga Jane o nagutta*
                    hit   I   hit
'John hit Mary, and then I hit Jane/John having hit Mary, I hit Jane.'   (Kuno 1973: 198)

(3c)
Taroo ga   hasit-ta   node, *boku wa Ziroo ni   mo soo sa-se-ta*
       NOM run-PAST   because I   TOP   AGT too so do-CAUS-PAST
'Because Taro ran, I made Jiro do so too.'   (Shibatani 1990: 316)

(3d)
aitsu   wa yuukan   da kedo, atama mo ii
that.guy TOP courageous COP but   head   also good
'That guy is courageous, and he is also smart.'   (Iwasaki 2002: 250)

Examples (3a)-(3d) show that subordinate clauses precede their main clauses.[7] Further, relative clauses come before their heads as in (3e), and complement clauses precede their matrix clauses as in (3f).

(3e)
[[biiru { ga/no } nomitai] hito]   wa dare desu ka
[[beer nom/gen drink:des] person] top who cop q
'Who are those people wanting to drink beer?'   (Iwasaki 2013: 213)

(3f)
maki   ni zehi   ai-tai   to (taroo wa michiko ni) itta
(name) dat really meet:des qt (name) top (name) dat say:past
Taro said to Michiko that (he) really wanted to see Maki.   (Iwasaki 2013: 220)

However, variations to these 'basic' patterns are frequently observed, especially in conversation, where clauses marked as subordinate may appear on their own, just as they do in Finnish.

        For the purposes of this paper, our definition of the clause conforms to that given in reference grammars of Finnish and Japanese. For Finnish, we take clauses to be syntactic units that contain a finite verb. For Japanese, clauses contain a final nominal, adjectival or verbal predicate; however, differently from the examples in 2 (a-e), Japanese clauses in our data often do not contain overt arguments of the predicate. Although Finnish and Japanese clauses and clause combinations are quite different, as we hope to have shown above, we assume for our purposes here, as argued by Thompson and Couper-Kuhlen (2005: 485), that the clause is a relevant unit for all languages, and specifically for Japanese and Finnish.[8]

---

        [6] The suffix *-te* connects temporally sequenced clauses.
        [7] Note that the distinction between subordination and coordination in Japanese is different from what is often found in European languages, in that what looks like a coordinator (e.g., *te* 'and/then' in 3b and *kedo* 'but' in 3d) may not be an independent lexical item but a suffix or clitic, making the clause look like a subordinate clause.
        [8] We are currently re-examining that claim (Laury, Suzuki & Ono 2013); our initial findings suggest that the relevance of the clause as a unit in the grammars of different languages is variable; minimally, it can be said that the clause is grammaticized to different degrees in different languages. However, our findings at this point are provisional.

4. Clause combining in Japanese and Finnish interaction

4.1. *Japanese*

When we turn our attention to conversational speech, we find something rather different from the clause combinations shown above. For instance, long sequences of utterances distinctly different from what we saw in (1)-(3) are commonly observed.[9] In what follows, we discuss examples from Japanese and Finnish to demonstrate this. In example (4), M is explaining how she and her friend were swaggering in the popular Shibuya district during Golden Week.[10]

(4) Swaggering

1	M:	*anoo are gooruden wiiku.*
		uhm well golden week
		'uhm well Golden Week'

| 2 | A: | *un.* |
| | | 'mhm' |

3	M:	*atashi hora konsaato gayoi shiteta janai.*
		I see concert attending doing:was TAG
		'see I was going to (lots of) concerts, right?'

| 4 | A: | *un=.* |
| | | 'mhm' |

5	M:	*sorede mata rei ni yotte sono senpai to*,
		and again as.usual that senior with
		'and again as usual with that older (friend)'

| 6 | X: | [*un*]. |
| | | 'mhm' |

7	M:	[*sono*] *nenpai yobawari sareta*[11] *sono sen*[*pai*] [*to*],
		that senior calling done:was that senior with
		'with that older (friend) who was called old'

| 8 | A: | [*un*]. |
| | | 'mhm' |

| 9 | H: | [*un*]. |
| | | 'mhm' |

---

[9] See also Ono (2011), who highlights the discrepancy between examples found in the literature dealing with syntax/grammar and examples found in conversation.

[10] Golden Week is a week-long holiday in Japan starting from the end of April and extending to the beginning of May.

[11] *sono nenpai yobawari sareta* 'who was called old' is a relative clause modifying the head *sono senpai* 'that older (friend)'. Generally there is no relative pronoun/relativizer in Japanese, as also shown above in 3e.

10 M: *futari de mata suupaa sapooto no anoo* [*pansuto*] *o haki*   [*no*],
              together again super   support of uhm   leggings O wearing FP
              'both dressed in Super Support leggings again'

11 H: [*haite*],
                                                            wearing
                                                            'dressed in'

12 A: [*un*].
                                                                              'mhm'

13 H: *un.*
              'mhm'

14 M: *minisukaato o haki*   [*no*],
              mini:skirt   O wearing FP
              'dressed in mini skirts'

15 H: [*un*].
                                  'mhm'

16 A: *un.*
              'mhm'

17 M: *soide anoo omoikiri   wakazukuri   o shite=*,
              and   uhm as.much.as.one.wants young:making O doing
              'and uhm dressed super young'

18 A: *un*.
              'mhm'

19 M: *de futari de mata shibuya no machi o ne*,
              and together again Shibuya of street O FP
              'and together again on the streets in Shibuya'

20 H: *un.*
              'mhm'

21 M: *ano nenuechikee hooru de konsaato ga at*[*ta   kara*],
              uhm NHK   Hall   at concert   GA existed so
              'since there was a concert in the NHK Hall'

22 A: [*un*].
                                                      'mhm'

23 M: *noshiaruiteta   wake*.
              swaggering:was FP
              '(we) were swaggering'

24 X: *un.*
              'mhm'

25 M: *sorede=*,
                and
                'and'

At first sight, the sequence in (4) involves a number of clauses. However, unlike the mono- and bi-clausal examples typically found in the literature and exemplified in (2) and (3), the predicates of many of these clauses appear without noun phrases which are required in the English translations (e.g., *noshiaruiteta wake* 'were swaggering' in line 23, which does not involve an expression equivalent to the English subject 'we'), a phenomenon often called ellipsis or zero anaphora.

   Besides these differences, what makes (4) look so different from (2)-(3) is its length. Here the sequence is transcribed based on intonation units. These units are either clauses (e.g., *soide anoo omoikiri wakazukuri o shite* 'and uhm dressed super young' in line 17) or phrases (e.g., *de futari de mata shibuya no machi o ne* 'and together again on the streets in Shibuya' in line 19). Several clauses in the sequence are found with dependent clause markers, which connect them to another clause. In fact, (4) is an example of so-called clause chaining (Iwasaki 1993b; Ohori 1992; Ono 1990), involving a series of dependent clauses used consecutively, as seen in *haki* 'wearing' in line 10, *haki* 'wearing' in line 14, *shite* 'doing' line 17, and *atta kara* 'since there was' in line 21, ending with the main clause involving a tensed verb form *noshiaruiteta* 'were swaggering' in line 23.[12] Similarly, noun phrases occurring independently in separate intonation units in (4) appear semantically or pragmatically tied to the predicates found in the sequence (e.g., *anoo are gooruden wiiku* 'uhm well Golden Week' in line 1 seems to go with *atashi hora konsaato gayoi shiteta janai* 'see I was going to (lots of) concerts, right?' in line 3 specifying the time period), further supporting the analysis that the entire sequence forms a unit.

   In fact, this long sequence appears to represent a discourse unit in which M starts to give a narrative about what her older friend and she did during the Golden Week break. It is interesting to note that the final clause *noshiaruiteta* 'were swaggering' ending this sequence in line 23 is followed by another particle *wake* 'the thing is', which has been shown to mark the end of mid-size discourse units in spoken language (Clancy 1980a; Suzuki 1991; on discourse units, see also Ford 2004 and references therein) in a way similar to paragraphs in written language. It should be pointed out that the above sequence is actually followed by another sequence of clauses (not shown in (4)) that begins with the conjunction *sorede* 'and' (see line 25). Intriguingly, this form has been shown to mark the beginning of a mid-size discourse unit in conversation (Sadler 2001). These observations thus suggest that the long sequence in (4) constitutes a mid-size discourse unit in spoken language, which ends with *wake* 'the thing is' and is followed by *sorede* 'and', starting another unit.

   It should be noted that clause chaining in Japanese is commonly observed in contexts of telling stories and making descriptions. In (4), M is the primary speaker (or the storyteller) and produces the whole sequence bit by bit, prosodically broken up into intonation units, represented by line breaks in the transcript. The production of these units seems to be coordinated with the brief verbal expression *un* 'mhm' (one instance

---

[12] Some predicates are found tensed and others not tensed in this example. Thus they might be labeled as finite and nonfinite respectively, based on the Western grammatical tradition. Note, however, that in Japanese, tense marking or the lack of it does not necessarily indicate main or subordinate clause status.

of so-called backchannel expressions in Japanese) produced by A and H (who might be called story recipients). The backchanneling expression *un* is found toward the very end (see lines 8, 9, 12, 15, and 22) or right after each clause/phrase that M produces (see lines 2, 4, 6, 13, 16, 18, 20, and 24).

As the production of each intonation unit by the storyteller becomes complete (or comes close to completion), a token of ratification is produced by the story recipients. That is, the content of each unit is negotiated involving such steps as the presentation of the story content by the former and its understanding and ratification by the latter; this collaborative process results in a combination of clauses. Each single unit projects continuation, and this projection is ratified by the recipient through the use of continuers.

Example (5), a conversation between two Japanese students concerning a recent earthquake, is again a fairly unexceptional sequence, yet it is perhaps an even more vivid example of how clauses are combined in Japanese talk, where the speaker constantly adjusts her talk to various factors that arise in the ongoing interaction. We will see that a proper analytic treatment of clause combinations requires that both verbal and non-verbal dimensions be examined, since the latter are found to be closely tied to the way clauses are combined in actual talk.

This sequence is divided into three parts (a), (b), and (c), not only to make it easier to see how the talk unfolds, but also to reflect relatively long pauses found at the beginning of (b) and (c);[13] long pauses have been suggested to correlate with higher-level breaks in discourse organization (Chafe 1987). We will now go over this example focusing on its prosodic, grammatical, semantic/pragmatic, and non-verbal aspects. Note that the transcription symbols for gaze and head nod are explained below.

(5) Jishin 1.2 (5:25)

1 (a) J: (h) *ano nazekatte iu   to anoo*,
                     uh   why   say if uh
                 'uh if (I may) say why uh'

                                          --
                                          H
2   *katagawa no ruumumeeto   ga*,
                one.side   of roommate   GA
                'the roommate on the other side'

3   K: H

                 -->
                                          H   H
4 J: .. *anoo detei-   muubuauto shita node*,
                     uh   went.out move.out   did   because
                 'uh had moved ou- had moved out so'

---

[13] These pauses are, however, not overtly indicated in the transcript because J takes inbreaths (marked by (h)) at the same time.

5 K: H   H   H   H   H

                                              H   H
6 (b) J: (h) *katagawa no anoo ookina ne hondana ga   aru deshoo.*
                      one.side of uh   big   FP bookcase GA exist COP
                  'uh there is the big bookcase on the other side, right'

7 K: H

                  ------>   ------------------------------------>
                                                            H
8 J: *a[no tsukue no ue   ni ookii hondana] ga arimasu nee,*
                  uh   desk   of top on big bookcase GA exist   FP
                  'uh (there) is a big bookcase above the desk, right'

                   HH H   H   HH
9 K: [*haa e   arimasu*].
                   oh   yes exist
                  'oh, yes, (there) is'

                                              H   H
10 (c) J: (h) *are ni hon ga haitte nakatta   mon de,*
                      that in book GA be.in wasn't   NML so
                  '(there) were no books in that so'

11 K: HH

                  --->
                        H   H
12 J: *are ga gataanto taoretan desu yo.*
                  that GA bang   fell   COP FP
                  'that fell (down with a) bang'

13 K: H

First of all, let us briefly describe what happens in the sequence. J is the primary speaker in this sequence; K gazes at J continuously throughout the sequence.[14] J explains why it took some time for her to realize there was an earthquake. J and K live in the same graduate dormitory and are familiar with the room layout. In (a) J starts giving her explanation by talking about her roommate's moving out of the shared housing. She appears to put this on hold after line 4, because in (b) she checks with K if he knows about the presence of the bookcase and its location. In (c), J goes back to the explanation that has been on hold since (a) by saying that the empty bookcase (due to the roommate's move and its being emptied) fell down. So looking at the entire sequence, it can be said that J provides K with an explanation in (a) and (c) following

---

[14] Since K's gaze at J is continuous throughout the sequence, it is not marked in the transcript.

the time sequence of the narrated events, while she interacts with K in (b) to confirm that K knows about the bookcase.[15]

Turning our attention to syntax, we note that example (5) includes a number of clauses which seem to exemplify the above-mentioned 'clause chaining': Several clauses in the sequence end with dependent clause markers such as *to* 'if', *node* 'because', and *de* 'so' (lines 1, 4, and 10 respectively). This indicates that these clauses do not stand on their own; they serve to project that a syntactic construction is still in progress and thus talk is continuing. The talk appears to be completed by *are ga gataanto taoretan desu yo* 'that fell (down with a) bang' in line 12, the final clause in the sequence, which takes a tensed verb and does not end with a dependent clause marker.

Interestingly, however, the sequence includes other clauses which are tensed and do not end with a dependent clause marker: *katagawa no anoo ookina ne hondana ga aru deshoo* 'uh there is the big bookcase on the other side, right' in line 6 and *ano tsukue no ue ni ookii hondana ga arimasu nee* 'uh (there) is a big bookcase above the desk, right' in line 8. These clauses in fact take the shape of independent clauses in traditional terms, though they appear in the middle of what could be analyzed as an ongoing syntactic unit, a clause chain. That is, J's explanation in (a) and (c) takes the form of a clause chain, containing a series of dependent clauses in lines 1, 4, and 10 and ending with an independent clause in line 12. The independent clauses in lines 6 and 8 are produced in (b) to see if K knows about the presence of the bookcase and its location. These clauses could be said to form a parenthesis (see, e.g. Duvallon & Routarinne 2005; Mazeland 2007).[16] They are syntactically separated from the rest of the sequence, and also constitute a separate action within the sequence.

Although the clauses in (5) might not globally form a syntactic unit such as a 'sentence' in traditional terms, their local production at each juncture of clauses is accomplished grammatically in the traditional sense while also responding to various on-line factors. That is, when we look at the individual clauses in (5) one at a time, we typically find clear one-on-one syntactic connections, such as the causal connection clearly marked between lines 10 and 12: *are ni hon ga haitte nakatta mon de* '(there) were no books in that so' and *are ga gataanto taoretan desu yo* 'that fell (down with a) bang' (See more about this below).

There are other indications that the sequence in (5) is connected as a unit. As noted earlier, the individual lines in (5), mostly representing clauses, correspond to intonation units, and we find that several of these units (lines 1, 2, 4, 8, and 10) end with 'continuing' intonation, represented by a comma in the transcript, projecting upcoming talk. The sequence ends with 'final' intonation, represented by a period, in line 12, which makes the utterance sound final; it does not project more talk in the same way 'continuing' intonation does, which supports the clause chaining structure of the sequence we suggested above. Obviously, there is another utterance, in line 6, that also ends with a final intonation, even though it occurs in the middle of the chained clauses. Again, this can be understood as a more local activity: J's attempt to confirm K's familiarity with the bookcase and its location, which arises in the middle of her more

---

[15] A similar but more syntactically based interpretation of this sequence is presented in Iwasaki & Ono (2002).

[16] As one reviewer rightly pointed out, utterances in lines 6 and 8 together with the response in line 9 can also be understood to form a side sequence (Jefferson 1972).

global activity of giving an explanation. Thus, overall we find that intonation mirrors the unit status of the sequence in (5).

Further, semantically, or perhaps pragmatically, many of the utterances represented in these intonation units are dependent on or connected to upcoming utterances. That is, the utterances *ano nazekatte iu to anoo* 'uh if (I may) say why uh' in line 1, *katagawa no ruumumeeto ga* 'the roommate on the other side' in line 2, *anoo detei- muubuauto shita node* 'uh had moved ou- had moved out so' in line 4, and *are ni hon ga haitte nakatta mon de* '(there) were no books in that so' in line 10 are not really complete on their own; instead they project further talk. That is, an explanation has been projected but it has not come to its end yet, as seen in the final form *de* 'so' in line 10. In fact, as we saw earlier, these utterances end with forms which are understood to syntactically project more material to come: *to* 'if' in line 1, *node* 'because' in line 4, and *de* 'so' in line 10 are dependent clause markers projecting another clause,[17] and in a similar manner, *ga* in line 2 is a nominal particle, often called nominative case (or subject) marker, projecting an associated predicate.

Finally, some of the clause chaining expressions found in (5) in fact project not just another clause but a discourse unit such as an explanation. For instance, J's clause *nazekatte iu to* 'if (I may) say why' in line 1, ending with the final dependent clause marker *to* 'if', projects that a syntactic construction is still in progress, yet it can be understood to further project an explanation for why it took some time for J to realize that there was an earthquake, and her talk is not really done until she completes her explanation. That is, the clause *nazekatte iu to* 'if (I may) say why' appears to have become conventionalized (Bybee's prefab; see Bybee & Hopper 2001) to project another clause which initiates further talk of a particular kind, in this case an explanation. A quick google search of the exact character string なぜかって言うと, which is just one of the several ways to write this expression in Japanese, resulted in 212,000 entries, showing that it is indeed a commonly used expression. Lexical items in this stock phrase seem to be fixed, and the word *naze* 'why' can be replaced with other WH words as illustrated below:

(6)

dooshite *ka tte iu to*	'if (I may) say why'
nande *ka tte iu to*	'if (I may) say why'
dare *ka tte iu to*	'if (I may) say who'
dore *ka tte iu to*	'if (I may) say which'
nani *ka tte iu to*	'if (I may) say what'
doko *ka tte iu to*	'if (I may) say where'
itsu *ka tte iu to*	'if (I may) say when'

This thus highlights another type of conventionalized linguistic resource that Japanese speakers use to project the upcoming talk, which further allows us to see the unit status of the sequence in (5).[18]

---

[17] An alternative gloss for *node* would be 'so', which would capture the positional characteristics of *node,* which appears at the end of the clause projecting another clause, similar to how English *so* functions.

[18] Much of actual talk seems to be produced using (semi-)fixed expressions. Fixedness appears to be a key feature in understanding the grammar of everyday talk (see, e.g., Hopper 1987). Obviously our description and theorization of grammar have to capture this aspect. Thus, considering that change is always ongoing, what we can come up with is at best an idealized synchronic grammar.

Looking at (5), one might be tempted to say that it represents a syntactic unit such as a sentence, but we have seen that the sequence is not grammatical in the traditional understanding. Further, the term 'unit' might suggest something fixed, but we saw above that much of the sequence is produced 'on the fly'; the particular combination of clauses included in the sequence and how they are combined as a whole is not preplanned (cf., e.g., Ford 2004). Thus, its final shape is adventitious and could not possibly be what J had planned to produce in advance. Rather it is an outcome of producing talk bit by bit, responding to various local cognitive and interactional needs, as, for example, observed in the confirmation-seeking sequence involving the two speakers in (b). This is also supported by the prosody; as we saw above, the sequence produced is broken up into intonation units, most of which are clauses, again giving the impression that J talks as she does her planning on-line. In addition, there are parts in the sequence that would not be considered semantically or pragmatically well formed in traditional grammar; this adds to our impression that it has not been preplanned. For example, J's utterance *anoo detei- muubuauto shita node* 'uh had moved ou- had moved out so' in line 4 ends with a dependent clause marker *node* 'because', which is expected to be followed by another clause. But what actually follows it, *katagawa no anoo ookina ne hondana ga aru deshoo* 'uh there is the big bookcase on the other side, right' in line 6, is not the clause one would expect (on semantic/pragmatic grounds).

Finally, the non-verbal aspects of (5) support our argument that the final shape of this sequence is not preplanned but is an outcome of talking bit by bit while responding to various local on-line factors including the interaction with K. As can be seen from the transcript, H marks one vertical head movement (hereafter, head nod); when head nods are produced along with speech by the same person, they are placed directly above the utterance line. Some head nods by K are produced without speech of his own; they seem to be produced as a response to the speech by J, the primary speaker in the sequence. To represent this relationship iconically, the head-nod marks are placed below J's speech with an empty line separating J's speech and K's head nods. The speed of the head nods varies considerably: Some are rather slow and others quick, often creating a sequence of repeated head nods. In addition, dotted lines with an arrow (----->), which are found above the head-nod line, mark J's gaze at K. K's gaze is not represented because she is gazing at J throughout the sequence. Given these non-verbal cues, we now observe that most of J's intonation units (mostly clauses) are accompanied by her own vertical head nods (see lines 2, 4, 6, 8, 10, and 12) toward the end. J's head nods begin with her first gaze at K (indicated by -->) found at the very end of line 2. They seem to be produced in order to have the intonation unit just produced ratified by K, an impression confirmed by K's response. That is, J's head nods in lines 2, 4, 6, 8, 10, and 12 are either immediately followed by, or co-occur with, head nods by K in lines 3, 5, 9, 11, and 13, as if K is responding to each of J's nods. Note that J's head nods at the end of line 6 are reciprocated by K's head nods at the beginning of line 9. This interpretation is supported by the semantic content and sequential nature of the two utterances forming the parenthesis, where J solicits K's ratification of the information about the bookcase by saying *katagawa no anoo ookina ne hondana ga aru deshoo* 'uh there is the big bookcase on the other side, right' in line 6. K in fact ratifies this by saying *haa e arimasu* 'oh, yes, (there) is' in line 9.[19]

_____

[19] There are only two occasions (lines 5 and 7) on which K's head nods are produced independently from J's head nods. It appears that at least in line 7 K is responding to J's recently

We have thus seen that the addition of gaze and head nod in our investigation enriches our understanding of the way clauses are combined in Japanese, demonstrating that a proper treatment of clause combining requires that both verbal and non-verbal dimensions be examined, since the latter are found to be closely tied to the way the former are produced. Specifically, we have seen that the head nods by the two speakers are tightly coordinated with the linguistic organization of the sequence, which, as we have shown, is being produced 'on the fly'.

4.2. *Finnish*

The next example comes from our Finnish data. It is an excerpt from the beginning of a videotaped conversation among three young women. Immediately prior to this, the participants have been drinking, and talking about, coffee. Miia then moves to a different topic. From what follows, it appears that she has taken a job in the town of Lahti for the next day, but is now wondering whether the weather is good enough to travel there. It appears that the other participants are already aware of Miia's plans from prior discussion. Each numbered line represents an intonation unit, but long intonation units have been divided into two lines for ease of glossing; in this case, the second line does not have a new number. Lines for gaze notations do not have numbers. Our discussion of this example is broken up into three parts to make the analysis easier to follow.

(7) SG151 Kaakkoissuomalaiset tytot_Kauppi_Myrsky

1   Miia: *itse asiassa*%,
                 self thing-INE
                 'in fact'

2   (0.5) *minä voin   soittaa sinne*,
                      1SG   be.able-1SG call   DEM3.LOC
                 'I can make a call there'

    Anu: -------

3   Miia: (0.6) *Lahtee   aamusta   et*,
                      Lahti-ILL morning-ELA COMP
                 'to Lahti in the morning'

    Anu: -----------------------------------

4   Miia: **jos** on   iha hirvee ilma   **et**
                 if   be.3SG quite awful weather COMP
                 '(to say) that if the weather is really terrible that'

    Anu: --

---

produced final particle *ne* (line 6), which has been shown to solicit collaboration among interactants (Cook 1992; Tanaka 2000).

Miia:	*hoitaa*	*jonku*	*muun*	*sinne,*
	arrange-3SG	someone-ACC	other-ACC	DEM3.LOC-LAT

'(they should) arrange (to have) someone else there,'

Anu: --> Miia

               --> Anu

5 Miia:	<X *ei* X> *mun*	*tarvii*	*täält*	*lähtee*	*ittee*[*ni--*]
	NEG 1SG-GEN	need	DEM1.LOC	leave-INF	self-PRT-1SGPOSS

'I don't have to leave from here (to V) myself'

Miia's turn in lines 1-5 can be said to constitute a clause combination, in which the three clauses in lines 1-4 can be said to be connected with clause linkers, as well as prosodically, while the last clause (line 5) is connected to the rest prosodically only. In terms of syntax, lines 1-3, *itse asiassa minä voin soittaa sinne Lahteen aamusta et* 'in fact I can (place a) call to Lahti in the morning (to say) that' could be analyzed as a matrix or 'framing' clause, which is followed by a clause combination functioning as its complement, *jos on iha hirvee ilma et hoitaa jonku muun sinne* 'if the weather's really awful that (they) arrange someone else (to come) there' in line 4. This complement, linked with *et* 'that' (end of line 3) to the preceding clause, consists of a *jos* 'if' clause, which, looking only at the connector, could in turn be analyzed as having its own complement, initiated with *et* 'so, that'. In that sense, the second clause in 4 could be seen as having at once features of a complement clause and of the main clause of the conditional clause that precedes it.

    This complex combination of clauses is produced smoothly, without trouble, although not in the fashion described in grammar books. First, the clause in lines 1-3 would not traditionally be analyzed as projecting a complement to follow; *soittaa* 'call' is not ordinarily considered a verb that takes a clause as a complement (see, for example, Hakulinen et al. 2004: 478-479, where *soittaa* is not mentioned in the list of verbs which appear in the referring construction).[20] However, it is not infrequently used that way. In Kirkkomäki's (2012) study of framing expressions of quotes in telephone conversations between three young women, *soittaa* was the third most frequently used verb before quotes (albeit well behind *olla* 'to be' and *sanoa* 'to say'; the latter is exemplified in 1b above), and in this use, it can certainly be analyzed as a complement-taking predicate (CTP; Thompson 2002) at least retrospectively. We might consider the use of a complement clause after *soittaa* an emergent pattern, and one that is conventionalized to the degree that the other participants recognize it. The analysis (or understanding) of *soittaa* as a CTP, and of the clause combination in which it appears as a matrix clause, emerges on-line as a consequence of the presence of the complementizer *et* at the end of the clause, which projects that what follows will be a complement. The grammar here (the syntax, semantics and pragmatics of the combination) is a temporal product, which is partly based on prior text, the combinations of matrix clauses ending with *et* and followed by complements, which the

---

[20] The list given in Hakulinen et al. (2004) is, of course, an open list. The grammar does note that verbs of communication go into this construction, even ones that are not ordinarily thought of as having quotes as complements. It is nevertheless interesting, given the frequency of *soittaa* in Kirkkomäki's (2012) corpus, that this verb does not appear in the forty-some verbs listed in the grammar. We therefore suspect that this use of *soittaa* is a relatively new one.

participants must have heard and used themselves previously, but the nature of this particular construction can only be understood as it unfolds.

The prosodic makeup of this segment also indicates that producing a complement after *soittaa* seemingly causes no extra planning effort, reinforcing the impression that the usage is commonplace. After *et* at the end of the clause, the complement follows smoothly, without any pause; there is a slight pitch reset after *et*, at *jos* (line 4), the only indicator that a new prosodic unit is starting.

Et is traditionally analyzed as a complementizer, but recent studies (Laury & Seppänen 2008; Koivisto, Laury & Seppänen 2011; Koivisto 2011) have shown that *et* functions not only as a complementizer, but also as both an initial and a final particle. In all three of its uses, it reorganizes the participant framework (Goffman 1981): It is used to index the introduction of another voice into the conversation (see also Hakulinen et al. 2004 on the use of *et* in the referring construction). It can introduce the voice of another participant, someone who is or is not a participant in the conversation, or the speaker's own voice earlier in the conversation or on another occasion. Here it marks what follows as a (prospective) quote, to be said by Miia to some others. The *jos* 'if' clause which follows *et* in line 4 could now be syntactically analyzed as a conditional clause within a complement clause, projecting a main clause to follow. However, what follows is another *et*-clause. What we might expect at this point on the basis of a frequency study cited by Hakulinen et al. (2004: 1068) is a *ni(in)* 'so' clause; the occurrence of *ni(in)*-clauses after initial adverbial clauses in clause combinations is described as "a rule rather than an exception" (2004: 1067; our translation).[21]

The relationship of the *et* clause in line 4 to the whole complex is problematic. It can be analyzed as a second complement to the matrix clause (as an (object) complement to *soittaa* 'to call'), marking what follows as a quote. In any case, the preceding *jos* clause does not qualify as its matrix clause, since it contains no CTP; yet there is a semantic connection between the *jos* clause and the *et* clause that follows it (a reasonable translation of what Miia says is, 'if the weather is terrible, they should arrange for someone else to come'). On the other hand, the *jos* clause could be considered parenthetical; Miia could be understood to be adding, after the first *et,* a condition which applies to her making a call, and then adding a complement clause, repeating the complementizer, which expresses the planned import of the call. However, what speaks against this analysis is that in this case, contrary to expectation, the parenthesis is not marked off prosodically from the rest of the utterance (cf. Duvallon & Routarinne 2005).

The last clause in this part, *ei mun tarvii täält lähtee itteeni* 'I don't need to leave from here (to X) myself', which could be intended as a justification for making the call to cancel the arrangements, is asyndetically linked to the preceding; there is no conjunction, and the preceding clause does not syntactically project continuation. However, the prosody at the end of the preceding clause may index continuation, as there is only a very slight fall at the end (line 4). In fact, the final clause of Miia's turn can be said to be linked to the preceding clause combination by a common line of pitch declination, if one compares relative pitch at the beginning of each intonation unit

---

[21] Hakulinen et al. (2004: 1068) note that a *ni(in)* clause follows an adverbial clause beginning with *jos* "nearly always" in spoken language: in the small study they cite, this happened in 63 of 65 cases.

(Couper-Kuhlen 2009; this volume), as the pitch at the beginning of line 4, at *jos*, is at 200 Hz, while the pitch at the beginning of line 5, at *mun*,[22] is at only 190 Hz.

What can we say about the grammar of clause combining here? The clause linkers as well as prosody are possible resources participants use to indicate that more talk is coming. The clause linkers mark boundaries between clauses and give clues about how the upcoming clause should be interpreted: As something said in another context (here, the future), that is, as a quote (*et* 'that'), or as a possible circumstance or state of affairs (*jos* 'if'). In other words, clause linkers not only link clauses to one another; they also, and perhaps even primarily, project the nature of the clause that follows them (Couper-Kuhlen 2012).

In terms of pitch and timing, the clauses in Miia's turn run smoothly together with no clear prosodic breaks. There is a slight pitch reset at *jos* 'if' (line 4), but at the end of the previous intonation unit, at *et* (end of line 3), there is only a slight fall, so even at this juncture, the prosody indexes continuation. The only two pauses are produced at maximally projective places, the first one at the end of line 1, after the discourse marker *itse asiassa* 'in fact', and the second one after the demonstrative *sinne* 'there(to)', line 2, which functions here as a kind of determiner of the oblique *Lahteen* 'to Lahti'. Further, sequences of clauses can be marked as continuing the same turn through a common line of pitch declination (lines 4-5). In other words, minimally, it is the clause linkers and the prosody that constitute the clause-linking 'grammar' here – they are what is shared in terms of projection by the participants in this conversation.

Clausal syntax, at least as traditionally described in grammars, seems to play a marginal role in the clause linkages here. The linkages appear to be planned one at a time, locally, and the whole sequence of clauses in Miia's turn does not hang together as a coherent whole. Even the transitions from one clause to another are not predictable from the nature of the preceding clause, including the argument structure of the predicate and the linkers used. For example, the verb *soittaa* may not by itself project a complement, although it can here be analyzed as an emergent CTP, especially when followed by *et,* and the *jos* clause can be seen as projective, but what follows it is not projectable from it.

The next part of the example is a direct continuation of the preceding. In line 5 (above), Miia has cut off her utterance when another speaker, Susa, starts speaking (see below); line 5 is incomplete because *itteeni* is in an object case (partitive), but the clause contains no verb which could take it as an object. The end of Miia's utterance here can be interpreted as reactive to what another participant does: She stops talking because someone else takes a turn. From the point of view of Susa, however, what are the possible grammatical cues that she is using to time her turn as she does? She might rely on what she knows about the argument structure of the verb *lähteä* 'leave': Its 'obligatory' arguments have already been filled at this point,[23] and in that sense, a projectable TRP has been created. This shows that turn continuation and turn transition, and hence the linking of additional clauses to a turn, are collaboratively determined. Miia's turn ends when Susa takes a turn, and thus the end of Miia's clause and the end of what we can retrospectively analyze as a clause combination can be seen as

---

[22] The pitch on *ei* may be even lower, but since this is an uncertain hearing, we have considered the pitch on the following word.

[23] The verb *lähteä* could be considered to project in addition to a subject or first argument (here, *mun* '1SG-GEN', which is genitive due to the deontic verb *tarvita* 'to need'), at least one oblique argument expressing the location of origin, here the locative demonstrative *täält* 'from here'.

responsive to Susa's action of taking a turn at the point she does. At the same time, we can say that Susa uses the syntactic makeup of the clause Miia is constructing to time the beginning of her utterance, that is, her timing is calibrated with what Miia has done.[24]

Anu has turned her head toward Miia at *et* in line 4, maintaining her gaze throughout Miia's turn, and seems to be vocalizing briefly (represented by # # #) in response to Miia's turn, in overlap with the beginning of Susa's turn in line 6. Until the beginning of line 5, Susa is standing between the camera and Miia, but at that point Miia becomes visible to the camera, and it can be seen that she is gazing at Anu as well. They break off their mutual gaze after the overlap. Susa's turn responds to Miia's preceding turn. The particle chain *no ei nii* 'well no' (line 7) expresses agreement with the last part of Miia's turn, *ei mun tarvii täält lähtee* 'I don't need to leave from here'. Then Susa recycles the idea of upcoming bad weather in line 9.

(8)

```
    Miia:        ---------> Anu

                 ---------> Miia
6   Anu:         [# # #]

7   Susa:        [no  ei    nii]  jos se    ois        [ni]mittäin--
                 PTC NEG PTC   if  DEM3 be-COND namely
                 'well no if (there) namely'

8   Miia:                                           [mm],

9   Susa:        jos se    on ainaki  sellanen myrsky niiku%
                 if   DEM3 be at.least such     storm  like
                 'if there's at least the kind of storm that'

                 (0.6) (h) ne      lupai[li],
                           DEM3 promise.FREQ-PST
                 'they were forecasting'

                                 -----------------------------> Susa
10  Miia:                        [tot]takai ne       sen        tajuaa
                                 of course DEM3.PL DEM3-ACC understand-3SG
                                 'of course they will understand it'

                 jos mie sanon    et,
                 if   1SG say-1SG COMP
                 'if I say that'

11  Susa:        (1.0) mm,

12  Miia         (0.3) joo et      emmie           lähe ttäält
                       PTC COMP NEG-1SG-1SG leave DEM1.LOC-ABL
                 'yeah that I am not leaving from here'
```

---

[24] An analysis of the overlap is provided by Vatanen (forthc.).

> *nyt   yhtää mihinkää*   ***et***, ((PUTS CRACKER IN MOUTH))
> now any   anywhere-ILL COMP
> '(to go) anywhere now that'

Susa begins a *jos* 'if' clause in line 7, and then produces a repair in line 9. These clauses are a partial repetition of what Miia has said in line 4. Together, they could be seen to confirm that Miia does not need to go to Lahti if the weather turns bad (line 7) and perhaps even that the people in Lahti whom Miia is planning to call should arrange for someone else to come.

Syntactically, line 9 might be analyzed as a *jos* 'if' clause, ***jos on ainaki s̲ellanen myrsky*** 'if there is the kind of storm', which has an embedded relative/adverbial clause *niinku ne lupaili* 'like they were forecasting', modifying it, but no main clause. On the other hand, the negative part of the particle chain *no ei niin* in line 7 could be analyzed as constituting the main clause for the *jos*-clause(s), since *ei* could be analyzed as a (repeat of) the negative verb. Susa's turn ends when Miia comes in in line 10, at the syntactically (though not prosodically) determined TRP at the end of Susa's clause. At this point, she turns her head toward Susa, who has moved to the left; Susa is not visible to the camera now.

The complex combination in lines 10 and 12 consists of what could be analyzed as a main clause containing a CTP *tottakai ne sen tajuaa* 'of course they will understand it', where the pronoun *sen* 'it-ACC' could be said to be either anaphoric, linking back to Miia's earlier turn (line 4) reporting what she plans to say, or cataphoric, projecting forward to the rest of the turn (lines 10 and 12). In any case this clause, possibly a complement-taking construction, is followed by a conditional *jos* 'if' clause *jos mie sanon et* 'if I say that' with an embedded *et*-complement *joo et enmie lähe täält nyt yhtää mihinkää et* 'yeah that I won't leave from here (to go) anywhere'. The verb of saying with *et* is a highly conventionalized expression and could be considered a prefab which strongly projects reported speech to follow. After a one-second pause and Susa's response particle *mm*, and a brief pause of 0.3 seconds, the projective reporting clause ending with *et* is followed by the complement in line 12.

It is worth noticing that here as well, the speakers pause (in line 9 and the beginning of line 11) after, and not before, clause linkers, not at what might be otherwise thought of as the end of the clause. These are, then, at least potentially projective places. In line 9, we have the linker *niinku* 'like' before a pause. The word ends with a glottal stop, and is followed by a half-second pause; such a pause after a conjunction can be analyzed as one designed to hold the turn (Local & Kelly 1986: 195; see also Laver 1994: 538). In addition, the phrase *semmonen* X *niinku* 'such an X that' also projects continuation. In line 10, we have the clause linker *et* before another pause. Although *et* can also be turn-final, after reporting phrases it normally projects more to come (albeit depending on sequential and actional factors; Koivisto 2011: 173-178). The clause linker *et* is repeated in line 11, after the pause and the particle *joo*. In these cases, the linkers, followed by a pause, and together with the clauses that precede them, project more to come and appear to function to hold the current speaker's turn.

The final *et* in line 12 is not followed by a complement clause – Susa's turn (see below) follows it smoothly. After the final *et*, Miia puts the cracker into her mouth. One possibility is that the final *et* may have been used here as an utterance/turn-final particle on purpose. Final *et* particles often occur after turns functioning as explanations that legitimize, paraphrase or explain something said before. As such, they can function

in spoken Finnish as signals of turn transition (Koivisto 2011). In other words, the final *et* can be seen as a signal that another speaker may now take a turn. On the other hand, it may just have ended up as such; in this case, the fact that she does not continue her turn can be seen as motivated either by the eating activity or by the turn Susa takes in line 14. That is, Miia may have intended to continue, but since Susa took a turn, she may have abandoned the plan, and yielded the turn. Alternatively, the *et* particle and the start of the eating activity may together work to signal the end of the turn.[25]

The last part of the example follows the previous one immediately. This part also shows that clauses can be combined only prosodically, without any clause linker (see Couper-Kuhlen 2009).

(9)

13 Susa: *siis   miks niil   ei   oo siel   tuntiohjaaja-a*,
              PTC why DEM3.PL-ADE NEG be DEM3.LOC-ADE hour-leader-PRT
              'so why don't they have a teacher there,'

14   *onks   ne   nyt kaikki sit jossai   vai*.[26]
             be-Q-CLT DEM3.PL now all   PTC somewhere or
             'are they then all somewhere (else) now.'

Simultaneously moving in from the left, into the space where the other participants are sitting, Susa asks a question *siis miks niil ei oo siel tuntiohjaajaa* 'so why don't they have a teacher there' (line 13) and then (line 14) provides a candidate answer formulated as a question, *onks ne nyt kaikki sit jossain vai*, 'are they then all somewhere (else) now'. The two clauses are asyndetically linked, although there is a slight rise at the end of line 13, and perhaps a slight pitch reset at the beginning of line 14. This turn does not continue; the turn-final *vai* is produced with creaky voice, which in Finnish can be taken as an indication that the turn is ending (Ogden 2004). After this excerpt, another participant produces a candidate understanding, and then Miia, who the question is addressed to, answers (not shown in the transcript).

What can we say about the grammar of clause combining here? What do the participants in conversation attend to? Projection can in fact be attributed to certain formulaic constructions: For example, *sellanen X niinku* 'such X as' (line 8) and *jos mie sanon et* 'if I say that' (line 9) do conventionally project continuation, and they are followed by more talk of a predictable type by the same participant in our data. On the other hand, although a type of clause thought to be projective is sometimes followed by another clause, what follows may not be the type of clause one would expect, as we saw in line 4, where the clause following a *jos* 'if' clause was a complement clause, even though the prior clause did not contain a CTP.

In fact, the actual combinations of clauses which appear in conversation do not neatly conform to the descriptions of clause combinations in grammars; instead, clause combinations seem to emerge locally, and are produced and processed on-line, in response to various factors, so that the end result is an adventitious outcome of producing clause-sized units and connecting them locally. For example, as we saw with the first clause in our example (7), what functions as a 'matrix' clause of a complement

---

[25] We thank Marja Etelämäki for this observation.
[26] Turn-final *vai* as a question particle is analyzed by Koivisto (in prep.).

clause may not contain a CTP recognized in grammars. Nevertheless, the verb *soittaa* 'call'*,* although not recognized as a complement-taking verb in grammars, does function, or emerges, as a CTP; that is, even when what is retrospectively analyzable as the matrix clause does not by itself project a complement, a complement may follow anyway. We also saw (line 4) that a complement clause initiated with *et* can follow a clause which contains no CTP, not even one retrospectively analyzable as such. Here, a complement clause followed nevertheless instead of a *ni* 'so' clause, which is thought to be the projected clause type after a *jos* clause in initial position. In this case, the *et* in line 4 could be thought to link syntactically and semantically backward to an earlier clause, the clause in lines 1-3, which might be retrospectively analyzable as its matrix clause, but the linkage would be across another clause (the *jos* clause in line 4). Moreover, the *et* clause in line 4 also has a semantic connection with the immediately preceding *jos* clause.

While clausal connections may be ambiguous, they do not seem to pose a problem for participants in conversation; Susa has no trouble responding to Miia's turn (lines 1-5) even though grammatically, the way the clauses are combined is not predictable from patterns familiar to linguists and described in reference grammars (see examples in 1), and the way the clauses are linked is, at the very least, syntactically ambiguous. Yet the response does not reflect any comprehension problems.

Argument structure on the level of the clause functions as a clue to participants as to when they can take a turn; for example, when Susa takes her turn (line 7), she takes it at a point where Miia's utterance (line 5) has come to a TRP in the sense that the projection of the main verb, *tarvii* 'need (to)', has already been fulfilled. At the same time, it seems that Miia might have been planning to continue her utterance (*itteeni* 'myself' is in an object case, but there is no verb yet which could take it as an object). However, Miia's utterance ends when Susa takes her turn.

Clause linkers in our data function as additional clues as to the type of clause that will follow, and to the degree that they occur at the edges of clauses, they can be said to mark clausal boundaries. They are projective in the sense that they occur at the end of a unit, but they do not always project turn continuation, since certain clause linkers (*ja* 'and'*, mutta* 'but' and *että* 'that, so (that)'; Koivisto 2011) can, and also do in our data, occur as final particles, as *et* does at the end of line 12.

Prosody also functions as an important cue to participants, marking projection in the sense of indexing the ends of clausal units, the beginning of new units, and the continuation of turns. Speakers place pauses strategically where the continuation of a turn has already been projected syntactically; at such places, there may be other prosodic signals which, like pauses, in another environment might function as indexes of turn yielding, such as glottalization (line 9). The syntactic projection is strong enough at these points, so that the other participants expect the turn to continue in spite of the pause and glottalization, and do not take a turn. On the other hand, new clausal units which have not been projected by what was said before them occur without any pause after the completion of the prior clause, and may be prosodically indexed as continuing the turn, as in lines 4-5, where a subsequent clause continued the pitch declination of the prior clause combination.

Finally, non-verbal activities such as bodily movement, gaze and activities the participants are engaged in also affect the way clause combinations are produced and when turns end and begin. For example, Miia's turn-final *et* (line 12) immediately precedes her putting a cracker in her mouth, and Susa's following turn (lines 13 & 14) is

coordinated with her movement back into the space where the other two participants are sitting. Miia has also addressed her turn to Susa by turning her gaze toward her in line 10. Susa's response particle in line 12 confirms this participation structure, as does her turn in lines 13-14.

5. Conclusions

Our examination of Finnish and Japanese conversations has revealed that clause combining in our data is managed locally and spontaneously, in response to cognitive and interactional factors that continuously change in interaction. The emerging structure is spontaneous, fluid, and contingent. The particular combination of clauses that comes out in the end is obviously not preplanned or memorized. Instead, parts are produced one at a time responding to various local, evolving factors. What might be incidentally said directly influences the actual shape of the upcoming talk, which is thus also incidental, and often not grammatical in the traditional view of syntax portrayed in the standard literature in linguistics (e.g., Stewart and Vaillette 2001; Fromkin and Rodman 1993; Payne 1997) or in the reference grammars of Finnish and Japanese cited above.

What, then, can we say might be the 'grammar' of clause combining? What do participants in interaction know about combining clauses? What they seem to know is how to combine clauses at a local level, in other words, how to combine the current clause with the next, or how to stop combining clauses, which is thus directly relevant to continuing or not continuing talk in interaction.

In both languages we examined, morphosyntax does provide a partial clue to how participants manage the production of clauses and, our main interest here, how they combine them. In both the Finnish and the Japanese data, argument structure appears to be used by participants in projecting whether a current turn will continue or is coming to an end. In both languages, morphological material at the end of clauses, especially conjunctions and other particles, are a key resource which provides the participants with clues to upcoming talk: Whether the talk will continue or not, and what the continuation is likely to be. However, these clues can be overridden. In both our Japanese and our Finnish data, what actually followed after a particle or conjunction was not always what might be conventionally thought of as projected.

Another resource used by participants in projecting the progress of turns is fixed clausal expressions or 'prefabs'. As we have seen above, they are used by speakers to project another clause that starts further talk of a particular kind. Much of what occurs in actual utterances is pieced together from 'prior text', formulaic bits of talk that have been experienced and used by speakers on many previous occasions. This seems to be a key feature in understanding everyday talk. These expressions may of course be more or less fixed at any point in time, and are a manifestation of the fact that linguistic resources comprise an ever changing, emergent set of partly formulaic, partly novel expressions.

Prosody is another resource for participants in constructing and detecting clause combinations. It is used as an index of turn continuation or turn ending, and prosodic patterns can also be used to show how, and whether, two clausal units are linked with each other. Participant responses are also sensitive to prosody. In particular, we saw above that the addressee in the Japanese conversation times his continuers to prosodically defined units of talk.

In addition, semantics and pragmatics also play a role in projection. Participants may wait to take their turn until a certain semantically or pragmatically projected action, such as an explanation, has been completed. We have also seen that non-verbal means, such as nodding in the Japanese data, and gaze direction in both data sets, are also finely attuned to and influence the structure of clause combinations and, more broadly, turn taking.

In our paper, we have discussed these different dimensions of verbal and non-verbal behavior that affect the ways clauses are combined in conversation as if they belonged to separate dimensions. However, in reality, they interact rather closely and our talking about them separately might simply be an outcome of biases in our training and more broadly of traditions in the field. Minimally, however, all of these dimensions must be considered to even begin to understand what is going on in combining clauses in actual talk. It seems obvious to us that these dimensions represent skills which Finnish and Japanese speakers have to combine clauses in actual talk; they are part of practice and knowledge. In that sense, we have now come to the most essential question for linguists: What should be considered as part of grammar? While it is clear to us that morphosyntactic factors alone cannot account for the patterns of clause combining in our data, it is not so clear what other dimensions of behavior should be included within the grammar of clause combining. Therefore, we are forced to leave this question open for now.

Looking at the on-line work of combining clauses in actual interaction microscopically but holistically, without being bound by the traditional method of examining a single factor at a time, has made it possible to identify and appreciate a variety of factors which are present when clauses are combined and relevant to how they are combined. Only through a close and more holistic inspection of actual interaction have we begun to see the relevant factors playing a role in the way clauses are combined in Finnish and Japanese.

We would also like to say that the findings we present here are still preliminary, in particular for the following reasons. The data we have examined are limited: Though we believe that some of the verbal and non-verbal behaviors of Finnish and Japanese interactants we have observed are rather typical, we have only been able to focus on a few excerpts. Obviously they do not encompass all genres or types of spoken Finnish and Japanese. To get a comprehensive picture of how speakers of Finnish and Japanese manage clause combining, we would need quantitative information from a large database involving a wider range of spoken discourse types. Our present results must, therefore, remain impressionistic, but we do hope that the disparate set of behaviors we have observed in connection with combining clauses in Finnish and Japanese will have provided us with some directions for future exploration on this topic.

Appendix

Gloss list
ABL   ablative
ACC   accusative

ADE   adessive
AGT   agent
CAUS   causative
COMP   complementizer
COND   conditional
COP   copula
DAT   dative
DEM   demonstrative
DES   desiderative
ELA   elative
FP   final particle
GEN   genitive
ILL   illative
IMP   imperative
INF   infinitive
LAT   lative
LOC   locative
NEG   negative
NML   nominalizer
NOM   nominative
PAST   past
PL   plural
POSS   possessive
PRT   partitive
PTC   particle
PST   past tense
Q   question
QT   quotative
REL   relativizer
SG   singular
TAG   tag question
TOP   topic
TRA   translative

Transcription symbol list

%   glottalization
-->   gaze
--> Anu   gaze direction
syllable (unhearable)
..   brief pause
(0.5)   timed pause in tenths of seconds
[talk]   talk within brackets is overlapped
<X talk X>   uncertain hearing

H	head nod
(h)	inhalation
?	rising intonation at end of utterance
.	falling intonation at end of utterance
,	continuing intonation at end of utterance
t<u>a</u>lk	strong emphasis
=	lengthening
ta-	interrupted word
(--)	unhearable utterance
---	interrupted utterance

References

Aho, E. (2010) *Spontaanin puheen prosodinen jaksottelu*. [Prosodic segmentation in spontaneous speech]. Ph.D. dissertation, University of Helsinki.

Auer, P. (1992) The neverending sentence: On rightward expansion in spoken syntax. In M. Kontra, and T. Váradi (eds.), *Studies in spoken Languages: English, German, Finno-Ugric*. Budapest: Hungarian Academy of Sciences, pp. 41-60.

Auer, P. (2005) Projection in interaction and projection in grammar. *Text* 25.1: 7-36.

Auer, P. (2009) Projection and minimalistic syntax in interaction. *Discourse Processes* 46.2: 180-205.

Auer, P., and S. Pfänder (2011) Constructions: Emerging or emergent? In P. Auer, and S. Pfänder (eds.), *Constructions: Emerging and emergent*. Berlin: Mouton de Gruyter, pp. 1-21.

Bybee, J. (2010) *Language, Usage and Cognition*. Cambridge: Cambridge University Press.

Bybee, J., and P. Hopper (2001) *Frequency and the Emergence of Linguistic Structure*. Amsterdam: John Benjamins Publishing Company.

Chafe, W.L. (1980) The deployment of consciousness in the production of a narrative. In W.L. Chafe (ed.), *The Pear stories: Cognitive, cultural, and linguistic aspects of narrative production*. Norwood: Ablex, pp. 9-50.

Chafe, W.L. (1987) Cognitive constraints on information flow. In R. Tomlin (ed.), *Coherence and grounding in discourse*. Amsterdam: John Benjamins Publishing Company, pp. 21-51.

Chafe, W.L. (1994) *Discourse, Consciousness, and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing*. Chicago: University of Chicago Press.

Clancy, P.M. (1980a) *The Acquisition of Narrative Discourse: A Study in Japanese*. Ph. D. dissertation, University of California, Berkeley.

Clancy, P.M. (1980b) Referential choice in English and Japanese narrative discourse. In W.L. Chafe (ed.), *The Pear Stories: Cognitive, cultural, and linguistic aspects of narrative production*. Norwood, NJ: Ablex, pp. 127-202.

Cook, H.M. (1992) Meanings of non-referential indexes: A case study of the Japanese sentence-final *ne*. *Text* 12.4: 507-539.

Couper-Kuhlen, E., and T. Ono. (2007) Turn continuation in cross-linguistic perspective. Special issue of *Pragmatics* 17.4.

Couper-Kuhlen, E. (2009) On combining clauses and actions in conversation. In *Virittäjä* 3.
 http://www.kotikielenseura.fi/virittaja/hakemistot/jutut/couper-kuhlen3_2009.pdf

Couper-Kuhlen, E. (2012) Turn continuation and clause combinations. *Discourse Processes* 49: 273-299.

Couper-Kuhlen, E., and S.A. Thompson (2008) On assessing situations and events in conversation: Extraposition and its relatives. *Discourse Studies* 10.4: 443-467.

Du Bois, J.W. (1987) The discourse basis of ergativity. *Language* 63: 805-855.

Du Bois, J.W., S. Schuetze-Coburn, D. Paolino, and S. Cumming (1993) Outline of discourse transcription. In J.A. Edwards, and M.D. Lampert (eds.), *Talking data: Transcription and coding methods for language research*. Hillsdale: Lawrence Erlbaum, pp. 45-89.

Duvallon, O., and S. Routarinne (2005) Parenthesis as a resource in the grammar of conversation. In A. Hakulinen, and M. Selting (eds.), *Syntax and Lexis in Conversation*. Amsterdam: John Benjamins Publishing Company, pp. 45-74.

Evans, N. (2007) Insubordination and its uses. In I. Nikolaeva (ed.), *Finiteness: Theoretical and Empirical Foundations*. New York: Oxford University Press, pp. 366-431.

Ford, C. (2004) Contingency and units in interaction. *Discourse Studies* 6: 27-52.

Fromkin, V., and R. Rodman (1993) *An Introduction to Language*. Fort Worth: Harcourt Brace College Publisher.

Goffman, Erving (1981) *Forms of Talk*. Philadelphia: University of Pennsylvania Press.

Goodwin, Charles (2003) The body in action. In J. Coupland, and R. Gwyn (eds.), *Discourse, the Body and Identity*. New York: Palgrave/Macmillan, pp. 19-42.

Haddington, P., L. Mondada, and M. Nevile (2013) *Interaction and Mobility: Language and the Body in Motion.* Berlin: Mouton de Gryuter.

Hakulinen, A., M. Vilkuna, R. Korhonen, V. Koivisto, T.R. Heinonen, and I. Alho (2004) *Iso suomen kielioppi.* [The Comprehensive Grammar of Finnish]. Helsinki: Suomalaisen Kirjallisuuden Seura.

Helasvuo, M.-L. (2001) Emerging syntax for interaction: Noun phrases and clauses as a syntactic resource for interaction. In E. Couper-Kuhlen, and M. Selting (eds.), *Studies in Interactional linguistics*. Amsterdam: John Benjamins Publishing Company, pp. 25–50.

Hauser, M.D., N. Chomsky, and W.T. Fitch (2002) The faculty of language: What is it, who has it, and how did it evolve? *Science* 298: 1569-1579.

Hinds, J. (1982) *Ellipsis in Japanese*. Carbondale: Linguistic Research.

Hopper, P. (1987) Emergent grammar. *Berkeley Linguistics Society* 13: 139-157.

Hopper, P., and S.A. Thompson (2008) Projectability and clause combining in interaction. In R. Laury (ed.), *Crosslinguistic studies of clause combining: The multifunctionality of conjunctions*. Amsterdam: John Benjamins Publishing Company, pp. 99-124.

Iwasaki, S. (1993a) The structure of the intonation unit in Japanese. In S. Choi (ed.), *Japanese and Korean linguistics, vol. 3*. Stanford: Center for the Study of Language and Information, pp. 39-53.

Iwasaki, S. (1993b) *Subjectivity in Grammar and Discourse: Theoretical Considerations and a Case Study of Japanese Spoken Discourse.* Amsterdam: John Benjamins Publishing Company.

Iwasaki, S. (2002) *Japanese.* Amsterdam: John Benjamins Publishing Company.

Iwasaki, S. (2013) *Japanese.* Revised edition. Amsterdam: John Benjamins Publishing Company.

Iwasaki, S., and T. Ono (2002) "Sentence" in spontaneous spoken Japanese discourse. In J. Bybee, and M. Noonan (eds.), *Complex sentences in grammar and discourse: Essays in honor of S.A. Thompson.* Amsterdam: John Benjamins Publishing Company, pp. 175-202.

Jefferson, G. (1972) Side sequences. In D. Sudnow (ed.), *Studies in social interaction.* New York: Free Press, pp. 294-338.

Keenan, E., and B. Comrie (1977) Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8: 63-99.

Keevallik, L. (2003) *From Interaction to Grammar: Estonian Finite Verb Forms in Conversation.* Uppsala: Acta Universitatis Upsaliensis.

Kirkkomäki, E. (2012) Johtoilmaukset ja referointi kertomuksissa. [Framing expressions and quotation in Stories]. MA thesis, University of Helsinki.

Koivisto, A. (2011) *Sanomattakin selvää? Ja, mutta ja että puheenvuoron lopussa.* [Goes without saying? *Ja, mutta* and *että* at the end of a turn at talk]. Ph.D. dissertation, University of Helsinki.

Koivisto, A. (in prep.) Online emergence of alternative questions in Finnish with the conjunction/particle *vai* 'or'. Submitted to R. Laury, M. Etelämäki and E. Couper-Kuhlen (eds.), *Combining Clauses and Actions in Interaction.*

Koivisto, A., R. Laury, and E-L. Seppänen (2011) Syntactic and actional characteristics of Finnish *että*-clauses. In R. Laury, and R. Suzuki (eds.), *Subordination in conversation: A cross-linguistic perspective.* Amsterdam: John Benjamins Publishing Company, pp. 69-102

Kuno, S. (1973) *The Structure of the Japanese Language.* Cambridge: MIT Press.

Kuroda, S-Y. (1965) *Generative Grammatical Studies in the Japanese Language.* Ph.D. dissertation, Massachusetts Institute of Technology.

Kärkkäinen, E. (2003) *Epistemic Stance in English Conversation: A Description of Its Interactional Functions, with a Focus on I Think.* Amsterdam: John Benjamins Publishing Company.

Kärkkäinen, E. (forthcoming) *I Thought It Was Very Interesting: Conversational Formats for Taking a Stance.* Oulu: Department of English, University of Oulu

Laury, R., and E-L. Seppänen (2008) Clause combining, interaction, evidentiality, participation structure, and the conjunction-particle continuum: The Finnish *että.* In R. Laury (ed.), *Crosslinguistic studies of clause combining: The multifunctionality of conjunctions, TSL 80.* Amsterdam: John Benjamins Publishing Company, pp. 153-178.

Laury, R., and T. Ono (2010) Recursion in conversation: What speakers of Finnish and Japanese know how to do. In H. van der Hulst (ed.), *Recursion and Human Language.* Studies in Generative Grammar 140. Berlin: Mouton de Gruyter: pp. 69-92.

Laury, R., R. Suzuki, and T. Ono (2013) Questioning the clause as a crosslinguistic unit in grammar and interaction. Paper given in the *Workshop on Linguistic and Interactional Units in Everyday Speech, International Conference on Cognitive Linguistics*, University of Alberta, Edmonton, June 21-22.

Laver, J. (1994) *Principles of Phonetics*. Cambridge: Cambridge University Press.

Local, J., and J. Kelly (1986) Projection and 'silences': Notes on phonetic and conversational structure. *Human Studies* 9: 185-204.

Mazeland, H. (2007) Parenthetical sequences. *Journal of Pragmatics* 39: 1816-1869.

Nariyama, S. (2003) *Ellipsis and reference tracking in Japanese.* Amsterdam: John Benjamins Publishing Company.

Newmeyer, F.J. (2003) Grammar is grammar and usage is usage. *Language* 79: 682-707.

Ogden, R.(2004) Non-modal voice quality and turn-taking in Finnish. In E. Couper-Kuhlen, and C.E. Ford (eds.), *Sound patterns in interaction. Cross-linguistic studies from conversation*. Amsterdam: John Benjamins Publishing Company, pp. 29-62.

Ohori, T. (1992) *Diachrony in Clause Linkage and Related Issues*. Ph. D. dissertation, University of California, Berkeley.

Okamoto, S. (1990) Potential Complements and Complementation: The Case of Japanese. Unpublished ms., California State University.

Ono, T. (1990) *te*, *i*, and *ru* clauses in Japanese recipes: A quantitative study. *Studies in Language* 14: 73-92.

Ono, T. (2011) The actual status of so-called particle ellipsis in Japanese: Evidence from conversation, acquisition, diachrony, and contact. In S. Rice, and J. Newman (eds.), *Experimental and empirical methods*. Chicago: CSLI, pp. 180-189.

Ono, T., and R. Suzuki (2011) Toward a new understanding of the so-called zero anaphora in Japanese everyday talk. Paper given in the panel 'Beyond Pro-Drop: The Pragmatics of Subject Ellipsis and Expression from a Cross-Linguistic Perspective' *IPrA 12,* Manchester, GB.

Ono, T., and S.A. Thompson (1997) Deconstructing 'zero anaphora' in Japanese. Proceedings of *the 23rd Annual Meeting of the Berkeley Linguistics Society*, pp. 481-491.

Payne, T. (1997) *Describing morphosyntax.* Cambridge: Cambridge University Press.

Pekarek Doehler, S. (2011) Clause-combining and the sequencing of actions: Projector constructions in French conversation. In R. Laury, and R. Suzuki (eds.), *Subordination in conversation: A crosslinguistic perspective*. Amsterdam: John Benjamins Publishing Company, pp. 103-148.

Sadler, M. (2001) Iconically motivated use of the Japanese discourse makers *sorede, nde,* and *de* in conversation. *Southwest Journal of Linguistics* 22.2: 143-161.

Scheibman, J. (2001) Local patterns of subjectivity in person and verb type in American English conversation. In J. Bybee, and P. Hopper (eds.), *Frequency and the emergence of linguistic structure.* Amsterdam: John Benjamins Publishing Company, pp. 61–89.

Shibatani, M. (1990) *The Languages of Japan*. Cambridge: Cambridge University Press.

Stewart Jr., T.W., and N. Vaillette (2001) *Language Files: Materials for an Introduction to Language and Linguistics, Eighth Edition*. Columbus: The Ohio State University Press.

Suzuki, R. (1991) A study of *wake* in Japanese narrative discourse. In M. Takeuchi et al. (eds.), *Kotoba no mozaiku: Okuda Natsuko meiyo kyooju koki kinen ronbunshuu.* [Mosaic of words: Papers in honor of Professor Emeritus Natsuko Okuda]. Tokyo: Mejiro Linguistic Society, pp. 124-137.

Tanaka, H. (2000) The particle *ne* as a turn-management device in Japanese conversation. *Journal of Pragmatics* 32.8: 1135-1176.

Thompson, S.A. (2002) 'Object complements' and conversation: Towards a realistic account. *Studies in Language* 26.1: 125-164.

Thompson, S.A., and E. Couper-Kuhlen (2005) The clause as a locus of grammar and interaction. *Discourse Studies* 7.4/5: 481-505. [Reprinted in *Language and Linguistics* 6.4: 807-837, 2005]

Tsujimura, N. (1996) *An Introduction to Japanese Linguistics*. Cambridge: Blackwell.

Vatanen, A. (forthc.) *Responding in Overlap. Turn onset timing and social action in Finnish and Estonian Conversation.* Ph.D. thesis, Department of Finnish, Finno-Ugric and Scandinavian Studies, University of Helsinki.