Sequence organization:
A universal infrastructure for social action[1]

Kobin H. Kendrick[a,*], Penelope Brown[b], Mark Dingemanse[c], Simeon Floyd[d],
Sonja Gipper[e], Kaoru Hayano[f], Elliott Hoey[g], Gertie Hoymann[b], Elizabeth
Manrique[h], Giovanni Rossi[i], Stephen C. Levinson[b]

[a]University of York, UK
[b]Max Planck Institute for Psycholinguistics, the Netherlands
[c]Radboud University, the Netherlands
[d]Universidad San Francisco de Quito, Ecuador
[e]University of Cologne, Germany
[f]Japan Women's University, Japan
[g]University of Basel, Switzerland
[h]University College London, UK
[i]University of California, Los Angeles, United States
[*]Corresponding author. Department of Language and Linguistic Science,
University of York, York, YO10 5DD, UK.

This article makes the case for the universality of the sequence organization observable in informal human conversational interaction. Using the descriptive schema developed by Schegloff (2007), we examine the major patterns of action-sequencing in a dozen nearly all unrelated languages. What we find is that these patterns are instantiated in very similar ways for the most part right down to the types of different action sequences. There are also some notably different cultural exploitations of the patterns, but the patterns themselves look strongly universal. Recent work in gestural communication in the great apes suggests that sequence organization may have been a crucial route into the development of language. Taken together with the fundamental role of this organization in language acquisition, sequential behavior of this kind seems to have both phylogenetic and ontogenetic priority, which probably puts substantial functional pressure on language form.

1 Introduction

It has sometimes been supposed that the grammar of languages is essentially universal (Chomsky 1981), whereas the principles of use are strikingly divergent (Hymes 1974). But to anyone familiar with informal conversation in different cultures, the reverse seems much more plausible (Levinson 2000:xiv, Evans & Levinson 2009), namely that languages differ fundamentally at every level of organization from sounds through grammar to meaning, whereas the basic organization of talk exchange in conversation is strikingly convergent across cultures. The plausibility of strong universals in conversational organization has been increased by the demonstration of precise parallels in turn-taking timing across unrelated cultures and different modalities (Stivers et al. 2009, De Vos et al. 2015) and exactly parallel procedures for interactive repair across languages (Dingemanse et al. 2015). If indeed such a universal infrastructure for language use can be established it would have important repercussions for understanding how languages are acquired, would establish the functional constraints on their form, and in promising a bridge to communicative

behavior in other primate species may help us understand the evolution of our extraordinary communication system (Levinson 2016).

In this article we turn to yet another domain where there appear to be strong universals of interactive language usage, namely the sequencing of social actions across neighboring turns. The initial observation, due to the founders of conversation analysis, is that actions often come in pairs (Schegloff & Sacks 1973). Certain actions by a first speaker therefore make relevant certain actions by a next one, so these interactional structures are built out of the turn-taking system on the one hand (Sacks et al. 1974) and social actions on the other (Levinson 2013a). Schegloff (2007) has developed a systematic description of parts of this system, along the lines of Figure 1.
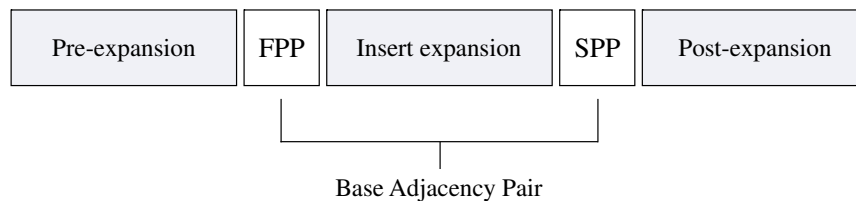
| Pre-expansion | FPP | Insert expansion | SPP | Post-expansion |

Base Adjacency Pair

Figure 1. The basic adjacency pair, with first pair-part (FPP) making relevant a specific second pair-part (SPP), and its optional expansions (Schegloff 2007).

The basic unit is the base adjacency pair, but this can be expanded by an additional pair before the base sequence, by the insertion of a pair between the base pair-parts, or by a pair after the base sequence is complete. Further, each of these slots can be recursively elaborated (Levinson 2013b), so building overall sequences of considerable length.

It is this organization that we shall argue in this article is essentially universal, demonstrating its operation across a sample of 12 languages and pointing out the functional motivations for each of its parts. It is prima facie clear that things could be otherwise; for example, actions could be interleaved so that first I produce a first pair-part of an adjacency pair (FPP) and then you produce another FPP, and then I respond to yours with a second pair-part (SPP) and you to mine with another SPP. (Indeed, something like this sometimes occurs in an academic lecture or press interview when the chair collects questions before they get answered.) It is the hypothesis of this article that the structure outlined in Figure 1, together with the possibility of recursive application of sequences, is universal in character. To be clear, the hypothesis is not that all interaction is made up of adjacency pairs and their expansions, but rather that such sequences occur in all languages as part of a universal infrastructure for social interaction (Schegloff 2006).

In this article we proceed as follows. First, we briefly describe the sample of languages in which we will test the hypothesis of the universality of the structure in Figure 1. Then we review briefly the underlying concepts, especially action, conditional relevance, normative expectations. Second, we start by demonstrating the cross-cultural validity of the base adjacency pair. Third, we explore pre-expansions, then fourth insert expansions, and fifth post-expansions and their exponence in the languages of our sample. Finally, we turn to consider the viability of the universality hypothesis given the data, and the language-specific features that we have found on the way. The article concludes with a discussion of the functional motivations of sequence organization, its phylogenetic roots in nonhuman primates, and its implications for the study of linguistic structure.

2   Data and method

To test the universality hypothesis, we drew on video corpora of informal social interaction between family and friends made in field sites across the globe. The sample included 12 languages, from 11 distinct language families, including both spoken languages and one sign language (see Table 1).

Language	Language family	Location	Researcher
ǂAkhoe Haiǁom	Khoisan	Namibia	Gertie Hoymann
Cha'palaa	Barbacoan	Ecuador	Simeon Floyd
English	IE (Germanic)	U.S. and U.K.	Kobin H. Kendrick
Italian	IE (Romance)	Italy	Giovanni Rossi
Japanese	Japonic	Japan	Kaoru Hayano
Argentine Sign Language	Italian Sign Language	Argentina	Elizabeth Manrique
Mandarin Chinese	Sinitic	Taiwan	Kobin H. Kendrick
Siwu	Kwa	Ghana	Mark Dingemanse
Turkmen	Turkic	Turkmenistan	Elliott Hoey
Tzeltal	Mayan	Mexico	Penelope Brown
Yélî Dnye	Isolate	Island Melanesia	Stephen C. Levinson
Yurakaré	Isolate	Bolivia	Sonja Gipper

Table 1. The languages in the sample and their respective field researchers.

The sample draws on all major continental areas, and includes languages with diverse subsistence patterns, from hunter-gatherer through swidden agriculture to post-industrial economies – thus from small-scale social systems through to complex stratified societies. In the sample, only English and Italian come from the same language family or cultural tradition. If we can find the hypothesized structures across these diverse languages, then a prima facie case has been made for their universality.

We began with the well-developed model of sequence organization based on English (Schegloff 2007), which proposes that the adjacency pair constitutes a basic interactional unit in conversation, while allowing for a set of methods for the expansion of adjacency pairs into more complex sequences of action as in Figure 1. Using methods of conversation analysis (see Clift 2016), we examined our corpora for evidence of adjacency pairs and six types of adjacency pair expansion: (i) generic pre-expansions, (ii) specific pre-expansions, (iii) post-first insert expansions, (iv) pre-second insert expansions, (v) minimal post-expansions, and (vi) non-minimal post-expansions (see Schegloff 2007). Definitions and examples are provided in Sections 3 and 4.

We adopted a qualitative approach and required clear evidence of the existence of each type of sequence in the relevant corpus to conclude that the pattern was attested in the language. At least three clear examples of each sequence type were thus required for each language, though for pre-second insert expansions the threshold was lowered to one example per language due to a relatively low frequency of occurrence. Candidate cases were submitted by each researcher to the first author who carefully vetted them and either accepted or rejected each as a clear example of a given sequence type. In general, our approach has been a conservative one: only clear and straightforward cases were accepted as adequate evidence for the existence of each sequence type. The supplemental materials include examples of all sequence types we could find for all 12 languages.

3   Fundamental concepts

Research on sequence organization began with a question. Given that an alternation between speakers is a basic feature of conversation, Schegloff (1968) asked, how can

one rigorously and empirically discriminate between a series of turns, in which one happens to follow the other, and a more complex structure, a sequence of turns, in which two turns cohere as a unit? The answer, Schegloff proposed, is that some turns at talk are united by a special property: *conditional relevance*.[2] If one turn is conditionally relevant on another, then the production of the first provides for the relevance of the second. A paradigm case is a summons-answer sequence, where the occurrence of a summons, such as calling out someone's name, makes an answer by the recipient conditionally relevant. If no answer occurs, its absence is a noticeable event, which allows for a variety of inferences (e.g. the recipient did not hear or is ignoring the summoner). Conditional relevance therefore sets up specific semiotic expectations – a response is due, even if it is not immediately forthcoming.

Conditional relevance is a defining property of an adjacency pair (Schegloff & Sacks 1973). Besides summons-answer sequences, other members of the class of sequences that constitute adjacency pairs include greeting-greeting, question-answer, offer-acceptance/refusal, request-granting/denial, invitation-acceptance/declination, compliment-acceptance/rejection, among others (Sacks et al. 1974:716).[3] An adjacency pair is defined as a unit of two turns by different speakers, normally adjacently placed one after the other, in which the first – a first pair-part – initiates an exchange and the second – a second pair-part – is responsive to the first and completes the sequence. The pair-parts are united by a relevance rule: upon the recognizable production and completion of a first pair-part, the recipient should produce a second pair-part of the same sequence type (Schegloff 2007:13-14 et passim).

The adjacency pair is a socially normative and socially accountable structure. The adjacency pair is not, as Heritage (1984b:246) put it, an "empirical generalization" about the frequency with which particular social actions occur in adjacent turns (see also Coulter 1983). Although quantitative studies have shown that first and second pair parts do indeed frequently co-occur, not only in English (Stivers & Robinson 2006, Kendrick & Torreira 2015) but also across a diverse sample of languages (Stivers et al. 2009), the adjacency pair is not, in the first instance, a statistical or probabilistic phenomenon. In the words of Heritage (1984b:247), "interaction is not structured, nor is its structure implemented, on the basis of statistical calculations". The adjacency pair should be understood as a form of normative organization: a first pair part establishes a set of normative constraints on the type and form of action with which the recipient should respond. The recipient may choose not to respond as expected – for example, by not returning a greeting – but this choice will nonetheless be understood as a meaningful departure from the norm.

Evidence for the operation of conditional relevance cited in the literature is fourfold. (i) The absence of a second pair-part after the recognizable production of a first pair-part is inferentially implicative, proving a warrant for inferences to account for the absence (Schegloff 1968), which can surface as sanctions (Stivers & Robinson 2006) or complaints (Atkinson & Drew 1979:54-55). (ii) If no second pair-part occurs,

[2] Antecedents to research on sequence organization appear in Sacks's lectures from the mid 1960s, most of which remained unpublished until 1992 (cf. Schegloff 1972:76). Sacks initially referred to adjacency pairs as "paired activities", using an exchange of greetings as an example (1965/1992a:95-99), and later as "paired objects", observing that "if a first member of a pair occurs, then the second ought to be done, and if it's not, that's noticeable" (1966/1992a:308). Sacks used this "paired-utterance phenomenon" to develop the notion of noticeable absences: "in the case of the pairs, where an utterance provides, specifically, the relevance of a second, the second's absence is noticeable and noticed" (1967/1992a:670). Schegloff (1967:86) notes that "the term and some elements of the idea of 'conditional relevance' were suggested by Harvey Sacks".

[3] Sacks (1967/1992a:667) provides "a non-exhaustive list" of actions that constitute first pair-parts: "greetings, challenges, threats, warnings, offers, requests, complaints, invitations, [and] announcements", to which he adds: "commands" and "demands".

the speaker of the first pair-part may pursue a response, using a variety of practices to do so (see Section 3.1). Pursuing a response reveals the speaker's expectation that a response was conditionally relevant (Atkinson & Drew 1979:52; Heritage 1984b:248-249). (iii) The recognizable production of a first pair-part provides for the relevance of subsidiary actions, such as repair initiations, that facilitate the completion of the sequence and thereby orient to the relevance of a second pair-part (Schegloff 1972, Atkinson & Drew 1979:55-57). (iv) A recipient who does not provide a conditionally relevant second pair-part, such as an answer to a question, may nonetheless orient to its relevance through the provision of an account for its absence (e.g. *I don't know*, see Heritage 1984b:249-251, Stivers & Robinson 2006).

3.1 The presence of conditional relevance across languages

The fact that questions, and by extension answers, are universally attested across the world's languages (Sadock & Zwicky 1985, Stivers et al. 2009) is prima facie evidence that the adjacency pair is a cross-linguistic universal. But evidence for the socially normative nature of conditional relevance has, to the best of our knowledge, only been adduced for familiar European languages (see e.g. Schegloff 2007 for English, and Stivers & Rossano 2010 for Italian). To determine whether the speakers of the languages in our sample orient to a normative obligation for a recipient to produce a second pair-part in response to a first, we examined sequences in which a first-pair-part speaker pursues a response from a recipient when it is not immediately forthcoming.

While we do not have the space here to exhibit the patterns across all dozen languages, the following case from Mandarin Chinese provides an example of such pursuits of response (see Section 1 in the Supplementary Materials for examples from all languages). As a mother and her adolescent son have an afternoon snack in the kitchen of their apartment, the son pours milk into a bowl of cereal and informs his mother that only a small amount remains. He then asks his mother to save the milk for him. Here and in the following extracts, we mark the first pair-parts with F and the second parts with S, indexing where necessary which parts belong together (see Appendix for annotation and transcription conventions).

(1) Mandarin Chinese (TPE09)

```
1   Son:   ni   kan  ta   zhi   sheng  yi   dian  dian  ei. ⌐
           2SG  look 3SG  only  remain one  point point PRT  │ F
           'look it only has a little bit left.'               │
                                                               │
2          (1.4)                                               │
                                                               │
3   Mom:   dui      a:.                                        │ S
           correct  PRT                                        │
           'yeah.'                                            ⌐┘

4          (1.0)

5   Son:   sheng  xia    dou gei  wo. ⌐
           remain down   all give 1SG  │
           'give me the rest.'         │
                                       │ F
6          yinwei weile     wo  yao   zhang gao. │
           because in.order.to 1SG want grow tall │
           'because I want to grow tall.'        ⌐┘

7          (1.2)                    ←— no response

8   Son:   bu   yao   gei  jiejie        le. ⌐
           NEG  want  give older.sister  PRT  │ F pursuit
           'don't give it to my sister.'     ⌐┘

9          (1.2)                    ←— no response

10  Son:   hao  bu   hao. ⌐
           good NEG  good  │ F pursuit
           'okay?'        ⌐┘

11  Mom:   °en-heng° (.)  hao   a. ⌐
           INTJ           good  PRT  │ S
           '°mm-hm° (.) okay.'      ⌐┘
```

The son's request for the milk (line 5), for which he provides a recipient-designed account (line 6), receives no response from his mother (line 7). At a position in which she could grant or deny the request, she remains silent. The son then issues an alternative request, namely not to give the milk to his sister (the only other person who lives in the apartment), an outcome that would likely increase his chances for the milk. Here, too, the mother remains silent (line 9). He then (line 10) uses a linguistic device more or less dedicated to pursuing a response in Mandarin Chinese: *hao bu hao,* literally 'good not good'. The disjunctive question explicitly solicits a response in which the mother either grants the request (e.g. with *hao* 'good') or denies it (e.g. with *bu hao* 'not good'). The mother grants the request at line 11, bringing the sequence to a close.

The pursuit of a response after a first pair-part provides evidence for the conditional relevance of a second pair-part. The absence of a second, such as the mother's silence in the example above, warrants the repetition or reformulation of the first pair-part, as well as the use of linguistic devices that explicitly pursue a response. In Tzeltal, a Mayan language of Mexico, there is a linguistic device that a third party – that is, neither the speaker of the first pair-part nor the recipient from whom a response is due – can use to intervene into a sequence (see Lerner 2019) to pursue a response. The expression *jak'a laj* is the imperative form of the verb *jak'* 'to answer' followed by the quotative particle, meaning that what the speaker has just said is a quote of what someone else has said. But the expression as a whole draws attention to the normative obligation to answer. A common use of this device is to pursue a response from a child.

In the following example, after Xn asks a four-year-old child to fetch a bowl, which is itself a pursuit after an earlier request, Mal intervenes on Xn's behalf, using *jak'a laj* to pursue a response.

(2) Tzeltal (Aluxosil)

1 Xn: ya bal a'-tzak-b-en   tatil?
            ICP Q   2E-take.in.hand-DIT-1A sonny $F_{pursuit}$
            'will you take it from me, little boy?'

2   ((pause))   ← no response

3 Mal: jak'-a   laj.
            answer-IMP QUOT $F_{pursuit}$
            '"answer", she says.'

4 Child: ehh
            INTJ   ← no response
            ehh

5 Al: suj-a   me a'-ba   tz'in.
            hurry-IMP PT   2E-REFL PT $F_{pursuit}$
            'so hurry up.'

6   ((lines omitted))

7 Al: yakuk. ya k-ich'   sch'aj.
            okay   ICP   1E-get-3A PT
            'okay. I'll get it then.'

Although the intervention fails to get a second pair-part from the child, it does elicit a vocal response in the form of a hesitation particle (line 4). The adults eventually abandon their efforts to secure a response from the child as Al volunteers to take away the bowl himself (line 7).

By pursuing a response, a speaker reveals his or her expectation that a response was due and treats its absence as an official and noticeable event in the conversation (Schegloff 1968, 2007). Response pursuits such as these were observed in all languages in the sample, providing evidence for the operation of conditional relevance in conversation across our diverse set of languages and cultures.

3.2 Not all adjacent actions are adjacency pairs: Broadcasting in ǂAkhoe Haiǁom

Before we proceed from adjacency pairs to their expansions, a crucial point must be made to delimit the scope of our investigation: conditional relevance is not the only principle by which sequences of action are organized. Participants use adjacency pairs to manage many basic social and communicative contingencies, such as transferring goods and services (offers and requests), exchanging information (announcements and questions), and initiating or terminating social encounters (openings and closings), among others (cf. Schegloff & Sacks 1973:297). However, not all sequences of action are organized as adjacency pairs (Schegloff 2007:9). In ǂAkhoe Haiǁom, speakers can produce turns at talk for others to hear with no obligation for them to respond. According to Hoymann (2016), such turns constitute 'broadcasts' similar to those described by Walsh (1991) for Australian Aboriginal languages. Broadcasts are produced at relatively low volume and do not employ practices, such as recipient-directed gaze, that address the turn to a particular recipient (Lerner 2003). In the following example, as three women sit outside and do beadwork, they use both adjacency pairs and broadcasts to organize their talk.

(3) ǂAkhoe Haiǁom (Ga Beads 2)

```
1   Ga:  ǀgui mûs tsū  e?    ((see Figure 1a))         ⌐        ⌐
          one  eye  only Q                            |        |  F
          'is it only one eye?'                       |        |

2         (0.2)                                        |        |

3   Su:  î                                             |        |
          yes                                          |        |  S
          'yes.'                                       ⌐        ⌐

4         (1.7)

5   Ga:  nēba nētse ra   dō-he      kara i   ge  ((see Figure 1b))
          DEM  today PROG bead-PASS  bead 3SN DECL           ← broadcast
          'here today beads are beaded.'

6         (0.8)

7   Ga:  Na  go !khom-de  ra   hau-ba-he       i   ge
          PSN PST bunch-3PN PROG bring-APPL-PASS 3SN DECL     ← broadcast
          'loads were brought for Na.'

8         (0.7)

9   Ma:  eso Na-s    ǀgôasa mâ sor  xū-e-s         go hau-ba-he    ⌐
          EXL PSN-3SF girl   Q  type thing-3SN-3SF PST bring-APPL-PASS  |  F
          'what type of thing was brought for Na?'                  |

10        (0.8)                                                     |

11  Ga:  kē    tama ta  ge  ((see Figure 1c))                       |
          look  NEG  1SG DECL                                       |  S
          'I don't know.' (lit. 'I didn't look.')                  ⌐
```


Figure 2. As she produces the broadcasts, Ga, in the middle of the frame, leans forward and lowers her head, attending to her beadwork, visible in panel (b).

The example begins with a polar question by Ga about the beadwork (line 1), which makes an answer conditionally relevant. After Su provides this, Ga leans forward, lowers her head, and returns to her work (see Figure 2b). Maintaining this position, she produces two utterances, neither of which is designed to solicit a response. She first formulates the activity in which the women are currently engaged: *nēba nētse ra dōhe kara i ge* 'here today beads are beaded' (line 5). Unlike her previous question, her formulation does not mandate a specific response from the other women, nor do they provide one. Indeed, like what Levinson (1988) has called an out-loud (see also Goffman 1981), the utterance lacks an addressee, being produced without gaze direction to either of the other women (cf. Stivers & Rossano 2010). Ga merely broadcasts the utterance to those around her as she attends to her work. After a short silence (line 6), she continues the broadcast with a report about the beads (line 7). While

this does elicit a next turn by Ma (line 9), the linguistic design of the turn is not one of a responding action (e.g. a response to a news announcement such as *abo* 'father!'), but rather that of an initiating action, a question. Ma requests specific information about the broadcast, namely the type of beads that were bought for Na. Crucially, Ga's broadcast did not call for a recipient to make such a request. The first pair-part of an adjacency pair, in contrast, specifies an action, from a delimited set, that a recipient should perform in the next turn. In general terms, broadcasts provide opportunities for broadcast-recipients to initiate sequences of action, but they do not establish normative obligations for them to do so.

Broadcasts in ǂAkhoe Haiǁom demonstrate that the socio-interactional principles that organize adjacency pairs are not the only principles that organize sequences of action in conversation. The relevance of a question after a broadcast is different in kind than the conditional relevance of a second pair-part on a first (see Schegloff 2007 on retro-sequences). In the remainder of this article, we focus exclusively on adjacency pairs and adjacency pair expansion across the languages of the sample.

4   Adjacency pair expansion

An adjacency pair can be expanded with additional adjacency pairs at three logical positions, each of which can be instantiated by two specific types, as shown in Table 2 (Schegloff, 2007). In this section, we illustrate the three positions and six types of sequence expansion that Schegloff describes for English, using data from our diverse sample of languages from around the world. A complete set of examples (i.e. of all types in all languages, where available) can be found in the Supplementary Materials.

Positions	Pre		Insert		Post	
Types	Generic	Specific	Post-first	Pre-second	Minimal	Non-minimal

Table 2. The three positions and six types of sequence expansion.

4.1 Pre-expansion

An adjacency pair that occurs before the first pair-part of a base adjacency pair and projects the relevance of that sequence is known as a pre-expansion. The nature of the action projection differentiates two types: (i) generic pre-expansions, which project the relevance of a base first pair-part but do not specify its action (e.g. a summons-answer sequence), and (ii) specific pre-expansions, which both project the relevance of a base first pair-part and specify or delimit the action it will perform (e.g. a pre-offer sequence).

Generic pre-expansions. The languages in the sample use a variety of linguistic forms to initiate summonses or generic pre-expansions, including particles (e.g. *hey* in English), names, honorifics (e.g. *me'tik* 'Mrs.' in Tzeltal), names with honorifics (e.g. *Odo san* 'Mr. Odo' in Japanese), kin terms (e.g. *papà* 'dad' in Italian), and person categories (e.g. *vieja* 'old woman' in Cha'palaa, borrowed from Spanish). In the following example from Mandarin Chinese, as Gu and Zhou talk about clothing, Wu interrupts with the particle *ei* 'hey', a summons addressed by gaze direction to Zhou.

(4) Mandarin Chinese (TPE15)

```
1  Gu:    you   de    pian[yi huo    you  hen  haoka[n.
              exist POSS cheap  goods also very good.looking
              'some bargain goods are also good looking.'

2  Wu:                    [ei.                                    ⌐
                          INTJ                                    |
                          'hey.'                                  | F_pre
                                                                  |
3  Zhou:                                      [dui    a.          |
                                              correct PRT         |
                                              'right.'            ⌐

4          (0.2) ((Zhou turns and gazes at Wu))                   |
                                                                  | S_pre
5  Zhou:  hah?                                                    |
          INTJ                                                    |
          'huh?'                                                  ⌐

6  Wu:    sorouyi. eh  w- (.) wo  nage::                          ⌐
          sorry    HES HES   1SG DEM                              | F_b
          'sorry. eh a- (.) I tha::t'                             |

7          (0.6)                                                  ⌐

8  Zhou:  dangan mei dai   lai    shi bu  shi.                    ⌐
          file       NEG carry come  COP NEG COP                  | F_ins
          'you forgot to bring the files, is that it?'            |

9  Wu:    dui. dangan jiu  shi mei dai   [lai.                    ⌐
          correct file    PRT COP NEG carry come                  | S_ins
          'yeah. I forgot to bring the files.'                    ⌐

10 Zhou:                                   [hou                   ⌐
                                           OK                     |
                                           'OK.'                  |
                                                                  | S_b
11         wo   ga  yi   giu  a     ((Taiwanese))                 |
           1SG  for 3SG  call PRT                                 |
           'I'll call him/her.'                                   ⌐
```

The particle *ei* 'hey' initiates a summons-answer sequence and thereby makes a response by the addressed recipient, Zhou, conditionally relevant. In response, Zhou directs her attention to Wu and answers with the particle *hah?*. This completes the pre-expansion sequence and allows for the initiation of the base sequence, an apology for a minor offense. The first pair-part of a generic pre-expansion such as this signals *that* a base sequence will occur, contingent on the response, but not *what* that sequence will involve. Generic pre-expansions such as this were observed in all languages in the sample (see Section 2.1 in the Supplementary Materials for examples). In Argentine Sign Language, where one needs to be assured of visual attention before signing, generic pre-expansions proliferate as signers frequently use a variety of visual and tactile actions – from waving, to tapping, to stomping on the ground – to secure the visual attention of a recipient before a base first pair-part (see Extract S18 for a striking example).

Specific pre-expansions. The first pair-part of a pre-expansion sequence can also specify or delimit the action that a projected first pair-part will perform. In the following English example, from an interaction between a group of university students in a common area on campus, Mark overhears Owen talk with other students about

video games. He then asks Owen if he has a video game console as a preliminary to a request to play together (i.e. a pre-request).

(5) English (RCE22a 41:42)

1	Mark:	Owen you have a console?	F_{pre}
2		(2.8) / ((Owen talks with others))	
3	Mark:	°Owen.° ((points to Owen))	F_{pre}
4		(0.5)	
5	Mark:	((raises and waves hand))	F_{pre}
6	Owen:	oh sorry.	S_{pre}
7	Mark:	do you actually have a console?	F_{pre}
8	Owen:	yea:h I have a PS3.	S_{pre}
9		(0.7)	
10	Mark:	you should bring it over sometime.=	F_b
11	Owen:	=I'll bring it over.	S_b

With no response to his question (line 1), Mark uses a series of generic pre's to secure Owen's attention. He first summons him by name, using a whisper voice that orients to his actions as interruptive (line 3), and then raises his hand into Owen's field of view (line 4). After Owen answers the summons, Mark reissues his question, one that ostensibly concerns Owen's possession of a video game console. Abstract questions such as this are commonly understood by participants not as requests for information in their own right, but as preliminaries to requests (Levinson 1983, Schegloff 2007). Such questions initiate specific pre-expansions, making conditionally relevant a response that either allows the projected base sequence to go forward, as Owen's response at line 8 does, or blocks it (e.g. "no, I don't"), an outcome that would allow the participants to avoid a rejection to the request proper. The request then follows in line 10. Specific pre-expansions were observed in 11 of the languages in the sample (see Section 2.2 in the Supplementary Materials for examples). A particular cultural practice in Tzeltal can result in a multitude of specific pre-expansions in a single sequence: An invitation to enter one's home can serve as a preliminary to an offer of food, which is obligatorily provided to guests. Yet guests routinely refuse multiple times before they accept such invitations, resulting in a proliferation of pre-offer sequences (see Extract S33).

4.2 Insert expansion

An adjacency pair that occurs between the first and second pair-parts of a base sequence and is initiated by the recipient of the base first pair-part is known as an insert expansion (1990, 2007:97-114). Two types can be differentiated according to whether they are backward or forward looking: (i) post-first inserts have a retrospective orientation and manage troubles in hearing or understanding a base first pair-part, and (ii) pre-second inserts have a prospective orientation and establish the conditions to decide between alternative base second pair-parts.

Post-first insert expansions. Because the first pair-part of an adjacency pair makes a response by the recipient conditionally relevant, it also requires that the recipient has heard and understood it well enough to respond appropriately. But this is not always

the case. The recipient may therefore produce an action that initiates repair by the first speaker. In the following example from Siwu, Ogbe asks Kuma a question about 'that one', an indexical reference to a person that Kuma evidently fails to recognize.

(6) Siwu (Two_men_2_698800; Dingemanse 2015:235)

1   Ogbe: nɛ kɔ̃rɔ̃ fɔɔ̀ sɔ   be   ǹgɔ́mmɔ̀ kà   ɔ̀   to   ɔ̀   bárá kɔ̃́rɔ̃̀: F_b
            so now you said what that.one IMM 3SG PROG 3SG do   now:Q
            'so now what did you say is that one doing right now?'

2   Kuma: ǹna: F_{ins}
            who:Q
            'who?'

3   Ogbe: ɛ̀ɛ̀ fɔ   ɔbi: (0.5) bɛ̃rɛ̃rɛ̃ĩ   nyɔarè gɔmmɔ̀, ɛ̀ɛ̀: S_{ins}
            HES your child   male.DIM second that.one HES
            'uh your uh (0.5) second born boy that one, uh:'

4   (0.8)

5   Kuma: ɛ̀ɛ̀ Kofi: F_{ins}
            HES PSN:Q
            'uh, Kofi?'

6   (0.6)

7   Ogbe: Kófi. Kòfi koso   ɛbɛrɛ [be   ũ S_{ins}
            PSN   PSN young or   what 3SG.TP
            'Kofi. Kofi junior or whoever'

8   Kuma: [Kòfi koso: F_{post}
                                  PSN   young
                                  'Kofi junior?'

9   Ogbe: ((head nod)) S_{post}

10 Kuma: ɔ   pia. S_b
            3SG be.located
            'he's around.'

In next position to the question, Kuma initiates an insert expansion, using an interrogative pronoun to locate the person reference in the questions as a source of trouble. Ogbe's response fails to resolve the trouble and Kuma initiates a second insert sequence, proffering a candidate understanding of whom 'that one' refers to, which Ogbe confirms (lines 5-7). The sequence that follows at lines 8-9 is not an insert expansion per se, as it targets not the base first pair-part but rather the second pair-part of the prior insert sequence (see Section 4.3 on post-expansion). After the proper referent has been established, Kuma produces the relevant second pair-part for the base sequence. Post-first insert expansions are sequences of other-initiated repair (Schegloff et al. 1977; Schegloff 2007), with which the recipient of a first pair-part addresses a trouble in hearing or understanding. Given that repair sequences of this type have been identified in all languages investigated to date (see e.g. the sample of 21 languages in Enfield et al. 2013), the universality of post-first insert expansions in the current sample is to be expected (see Section 3.1 in the Supplementary Materials for examples).

Pre-second insert expansions. Although some first pair-parts require that the recipient produce a second pair-part of the same action type (e.g. a greeting calls for a reciprocal greeting), many allow the recipient to select among a set of alternative second pair-parts (e.g. one may accept or decline an invitation). To facilitate this selection, the recipient can initiate a sequence whereby he or she solicits additional information from

the speaker of the first pair-part. Such sequences do not retrospectively address troubles in hearing or understanding a base first pair-part but rather prospectively establish the conditions for the base second pair-part. The following example of a pre-second insert expansion comes from a northern dialect of Italian as Bianca and Flavia take their turn in a game of cards.

(7) Italian (Circolo01)

1	Bianca:	meti		zo	'n	altro	ambo	
		put-NPST-2s	down	one	other	double		F_b
		'put down another double'						

2		(2.5) ((Flavia looks at her cards))

3	Flavia:	de sete	'l	g'ho	
		of seven	3s.A	EX=have-1s	
		'I have one of sevens'			

4		(0.4)	F_{ins}

5	Flavia:	o de quatro
		or of four
		'or of fours?'

6		(3.2)

7	Flavia:	de sete	o	de quatro	
		of seven	or	of four	F_{ins}
		'of sevens or of fours?'			

8	Bianca:	de quel che		te	voi	
		of that which	SCL	want-2s		S_{ins}
		'of what you want'				

9		(1.4)

10	Flavia:	meto	'l	sete	
		put-1s	the	seven	F_{ins}
		'(shall) I put sevens?'			

11	Bianca:	sì	dai	
		yes	PRT	S_{ins}
		'yes alright'		

12	Flavia:	((lays down a double of sevens))	S_b

After Bianca directs Flavia, her teammate, to put down a double (i.e., two cards of the same number), Flavia does not immediately produce the relevant second pair-part but rather offers Bianca two alternatives and thereby initiates a pre-second insert expansion. When Bianca selects neither, Flavia pursues the matter, initiating another insert expansion, a proposal of one alternative. Only after Bianca agrees does Flavia produce the conditionally relevant response. Here, then, the recipient of a first pair-part initiates two sequences designed to facilitate the selection a second pair-part from among a set of alternatives. Across the sample of languages pre-second insert expansions such as this were impressionistically quite rare, and indeed have thus far been observed in only nine of the 12 languages (see Section 3.2 in the Supplementary Materials for examples).

4.3 Post-expansion

Once the recipient of the first pair-part produces the conditionally relevant second pair-part, the adjacency pair is complete. But the completion of a base adjacency pair is not necessarily the end of the sequence. The speaker of the base first pair-part can expand the sequence in two basic ways: (i) through the production of a turn that registers the adequacy of the base second pair-part and explicitly signals sequence completion, or (ii) through the initiation of a sequence that takes the opposite stance towards the base second pair-part, indicating that it is especially newsworthy, or alternatively inadequate (and hence to be elaborated) or somehow problematic (Schegloff 2007:115-168). There are thus two types of post-expansion: minimal post-expansions, also known as sequence-closing thirds (SCT), and non-minimal post-expansions, which are themselves adjacency pairs. In comparison to expansions at other positions, post-expansions, especially non-minimal ones, exhibit greater variation across the languages in the sample.

Minimal post-expansions. After a second pair-part, the speaker of the first pair-part can produce an action that registers the second and signals that the sequence is complete. Such sequence-closing thirds are optional expansions of the sequence.[4] In the following example from Turkmen, after Myrat answers a question about whether a mutual friend owns a hookah, Maral registers the answer firstly with a particle *hä*, which functions similarly to the change-of-state *oh* in English (Heritage 1984a), and secondly with a claim of understanding.

(8) Turkmen (Gepleshligi 42:45)

```
1   Maral:  öz   öý-ün-de-de              bar   diý-di-däl-mi?          ⌉
            self house-3SG.POSS-LOC-EMPH  exist say-3SG.PST-not-Q      |
            'didn't he say he had one at his house?'                   |
                                                                       |  Fb
2           ýa  ýok-my.                                                |
            or  no-Q                                                   |
            'or no.'                                                   |
                                                                       |
3   Myrat:  maňa    aýd-anok-da,                                       |
            1SG.DAT tell-3SG.NEG.PRS.PRF-EMPH                          |
            'he hasn't told me anything,'                              |
                                                                       |  Sb
4           sor-ap    gör-emok.                                        |
            ask-CONV  see-1SG.NEG.PRS.PRF                              |
            'and I haven't asked.'                                     ⌋

5   Maral:  hä:. düşn-ük-li.                                           ⌉
            oh   understand-NMLZ-ASSO                                     SCT
            'oh:. I see.'                                              ⌋

7   Myrat:  Maýa bilen habar-laş-ýarsyň-my?                            ⌉
            PSN    with  news-COOP-2SG.PRES-Q                             Fb
            'do you keep in touch with Maya?'                          ⌋
```

Such actions in the third position of a sequence register the second pair-part and treat the adjacency pair as closed. In this example, the initiation of a new sequence by Myrat at line 7 after the sequence-closing third demonstrates that he has understood the prior

---

4 Jefferson and Schenkein (1978) have argued that some sequence types require a third position action to close the sequence (see also Kevoe-Feldman & Robinson 2012).

sequence to be complete. Minimal post-expansions were observed in all languages in the sample (see Section 4.1 in the Supplementary Materials for examples). In Argentine Sign Language, the body position of the one who produces the first pair-part often reflects the status of the sequence. A signer who asks a question may lean forward towards the recipient and hold this position until an answer is given. Then, in third position, the signer returns to his or her body to neutral position and thereby visibly embodies the closure of the sequence, which can be understood as non-verbal sequence-closing third (see Extract S63).

Non-minimal post-expansions. Four types of non-minimal post-expansion have been described by Schegloff (2007) for English: (i) topicalizing a second pair-part, thereby treating it as worthy of further on-topic talk; (ii) initiating repair on a second-pair-part; (iii) rejecting, challenging, or disagreeing with a second pair-part; and (iv) reissuing or reworking a first pair-part after a non-aligning response (p. 149-168). Aside from the first, topicalization, the non-minimal post-expansions identified in English are all remedial or disaffiliative in nature, used to manage various misunderstandings and misalignments that emerge with the production of a second pair-part.[5] This third position after the adjacency pair is a natural locus for such actions, since the second part of the base pair makes clear how the first was understood and acted upon (Schegloff, 1992).

But in two of the languages in our sample, Yurakaré and Tzeltal, an additional type of non-minimal post-expansion is especially common, namely repetitional post-expansions in which the speaker of the first pair-part repeats all or part of the second pair-part and in so doing makes confirmation by the speaker of the base second pair-part conditionally relevant – a type of non-minimal post-expansion we refer to as a *repetitional receipt sequence*. The repetitional receipt sequences in Tzeltal appear with an abundance unmatched by the other languages in the sample. The extreme proliferation of repetition in Tzeltal has been previously documented by Brown (1998, Brown & Levinson 1987) and appears in other Mayan languages as well. In the following example, two base adjacency pairs are expanded by repetitional receipt sequences, the second of which includes a sequence-closing third. The questions by Chanit concern Alonso's ex-wife and circumstances that lead to his time in jail.

---

[5] There thus appears to be a general tendency in English for sequence expansion – at all positions – to manage social disaffiliation and interactional misalignment.

(9) Tzeltal (0268dig1Chanit)

1 Chanit: jal   tz'i alal-etik i (.)
           but PT   child-PL PT
           'as for the children' (.)

2   y-ik'-tikla-ye   ya'tik tz'i mak a'w-ik'-oj-la   bal.
           3E-fetch-DISTR-PERF now   PT   PT   2E-fetch-PERF-DISTR Q
           'has she abandoned them now or have you taken them?'

3 Alonso: jo'o: y-ik'-tikla.
           no   3E-fetch-DISTR
           'no she abandoned them.'

4 Chanit: yik'tikla[:.
           3E-FETCH-DISTR
           'she abandoned them.'

5 Alonso: [y-i [k-
                    3E-fetch
                    'she aban-'

6 Chanit: [ja' niwan   ya chuk-wan yu'un   tz'i bi ma.
                            it's perhaps INC jail-ANTIP because PT PT PT
                            'maybe that's why she jailed (me) perhaps.'

7 Alonso: ja'   laj::
           it.is.so QUOT
           'it is, she says.'

8 Chanit: ja'   kati?
           it.is.so PT
           'it is so, golly.'

9 Alonso: °ja'   laj°
           it.is.so QUOT
           'it is, she says.'

10 Chanit: w-a'↑y:.
           1SG/2SG-see
           'I see.'

F_b

S_b

F_post

S_post

F_b

S_b

F_post

S_post

SCT

In contrast to a language like English, in which responses to polar questions normally employ polar response particles such as *yes* and *no* (Raymond 2003), in Tzeltal an affirmative response is normally accomplished by repetition (Brown 2010, Enfield et al. 2019). As a result, repetitional receipt sequences in Tzeltal typically consist of two repetitions, one as a receipt of the base second pair-part and one as a confirmation of the receipt. Two repetitional receipt sequences occur in the example above. After Alonso answers Chanit's first question, Chanit repeats the answer at line 4, which Alonso then confirms in the next turn, though he does not complete his utterance as Chanit issues a follow-up question in overlap at line 6. Similarly, the repetitional receipt by Chanit at line 8 also elicits a confirmation by Alonso, after which Chanit produces a claim of understanding and thereby brings the sequence to a close.[6] Unlike in the English cases described by Schegloff (2007), these non-minimal post-expansions are not disaffiliative in nature: the speaker who initiates the post-expansion does not challenge the base second pair-part, nor does he treat it as a source of trouble in need of repair (cf. Schegloff 2007:149-168). Repetitional receipt sequences in Tzeltal are a

---

6 The two part receipt is a very common format, but odd numbers of repeats occur, so it is clear that not all have the strict conditional relevance typical of adjacency pairs.

normal, unmarked way to register a base second pair-part and to bring the sequence to a close. While repetitional receipt sequences were especially prominent in Tzeltal, as well as Yurakaré (see Extract S81), non-minimal post-expansions more generally were observed in all languages in the sample (see Section 4.2 in the Supplementary Materials for examples).

4.4 Expansion types across languages

The primary goal of our investigation was to determine whether adjacency pair expansion, which has been well documented in English, also occurs in a diverse sample of languages and cultures from around the world. We have relied for exposition on select examples in this article and its supplementary materials. Our results show that adjacency pair expansion is indeed a cross-linguistic and cross-cultural phenomenon. We found that sequence expansion occurred in all three positions – pre, insert, and post – in all languages in the sample. Moreover, four of the six types were also well attested in all languages (see Table 3). However, two types of adjacency pair expansion were not universally attested: specific pre-expansions in Cha'palaa and pre-second insert expansions in Argentine Sign Language, Tzeltal, and Yélî Dnye. Furthermore, even in other languages, pre-second insert expansions were impressionistically rare and for some only a single clear case could be identified. Thus while we found unequivocal evidence of sequence expansion in all three positions in all 12 languages, particular types of expansion in pre and insert position were not universally attested.

Language	Pre		Insert		Post	
	Generic	Specific	Post-first	Pre-second	Minimal	Non-minimal
ǂAkhoe Haiǁom	●	●	●	●	●	●
Cha'palaa	●	-	●	●	●	●
English	●	●	●	●	●	●
Italian	●	●	●	●	●	●
Japanese	●	●	●	●	●	●
LSA	●	●	●	-	●	●
Mandarin	●	●	●	●	●	●
Siwu	●	●	●	●	●	●
Turkmen	●	●	●	●	●	●
Tzeltal	●	●	●	-	●	●
Yélî Dnye	●	●	●	-	●	●
Yurakaré	●	●	●	●	●	●

Table 3. The attestation of positions and types of adjacency pair expansion across the sample of languages. ● = attested, - = unattested so far.

Within the context of uniformity, however, we have also observed subtle variation in some languages. The visual modality of Argentine Sign Language, we have noted, places a special burden on generic pre-expansions, and such sequences thus occur with an abundance unmatched by the spoken languages in our sample. The cultural practice in Tzeltal of refusing pre-offers can lead to multiple pre-sequences being produced before a base adjacency pair. Such sequences, while characteristic of a particular cultural practice, are nonetheless built out of basic structures that occur across the languages in the sample. The repetitional receipt sequences found in Tzeltal and Yurakaré are similarly characteristic of these languages, but again the basic structures are the same as those found in other languages. The data thus suggest that the basic structures of sequence expansion in pre, insert, and post positions are remarkably uniform, even if particular uses that speakers of different languages find for them can differ in subtle ways. We now turn to consider what might lie behind these striking parallels in language use across cultures.

5   Discussion

5.1 The functional organization of adjacency pair expansion

Sequence expansion in human interaction can be shown to have some straightforward functional motivations. First, the adjacency pair is a basic unit of collaborative action. It establishes a division of labor between initiation and response as a cooperative solution to the problem of joint action (see Clark 1996). As such, it is a form of contingent action distributed across two parties, one whose roots can be seen in human ontogeny, where infants learn that smiles and vocalizations elicit stimulating and matching reciprocation from the caregiver (Bruner 1975, Hilbrink et al. 2015). The infant must go on to learn that different initiating actions get tailored responses, so that, for example, lifting the arms elicits being picked up by the caregiver. This kind of proto-adjacency pair can be learned through ritualization, so that what was at first a natural adaptation to being lifted becomes a signal of wanting to be lifted, and it can be observed in our nearest cousins the great apes especially between mothers and infants (see Section 5.3 below). It relies of course on a desire to help, and thus in humans it points to a fundamental generalization of co-operation beyond immediate kin (Tomasello 2008). Conditional relevance has its root then in the obligation to play one's part in joint action, to which human communication is assimilated (Clark 1996).

The core unit, the adjacency pair, is thus highly motivated. But what about the specific structures of pre-expansions, insert sequences, and post expansions that are built around it? In part following Schegloff (2007), our argument is that these can be seen to arise from the way in which the interactional production of a basic adjacency pair sequences generates recurrent contingencies that participants must manage (see Figure 3). The interactional system has basic requirements (Goffman 1976), akin to Hockett's (1960) design features of language. The interactional system requires mechanisms to secure an open channel, to implement a course of action jointly, to repair troubles if they emerge, and so on. Let us take each of the three expansion types in turn.

Consider first the functional motivations for pre-expansions. Generic pre-expansions or summonses are motivated by the fact that mutual attention (or channel accessibility) must be established before the base sequence is delivered, lest the recipient fail to see or hear the base first pair-part. Specific pre-expansions project the nature of the forthcoming base sequence, usually by checking that a precondition for the base FPP holds (as in "do you actually have a console?" in Extract 6). It is notable that in English at least specific pre's are used especially where the base FPP is a socially delicate action, such as the delivery of bad news, or a request or complaint that may threaten the recipient's face. In general, then, a preliminary sequence can pre-manage issues of face and dispreference by giving early warning about the nature of the base FPP, which can be aborted early if necessary (see Levinson 1983, Schegloff 2007).
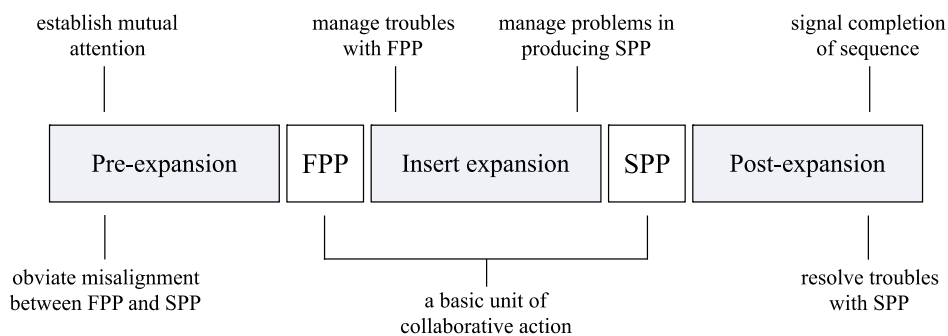
Figure 3. The functional organization of adjacency pair expansion.

Turning now to insert expansions between the base FPP and SPP, there are different functional motivations for the two types described above. Post-first inserts have as central cases repair sequences, elicited by an other-initiation of repair. Clearly the recipient of the FPP is in no place to produce a SPP if it was not heard or not understood. Pre-second inserts on the other hand are used when further information is required in order to select among alternative second pair-parts (see Extract 8). Insert expansions thus handle preconditions for a successful SPP that were missing in the delivery of the FPP.

Post-expansions are of two rather different types: single turns (sequence-closing thirds) and adjacency pairs. A single turn may register the adequacy of the SPP and thereby forego an opportunity to expand the sequence, indicating sequence closure (Heritage 1984, Schegloff 2007). One of the culturally divergent findings of this study is that in some languages this sequence-closing function is performed by a full adjacency pair (cf. the discussion of Tzeltal and Yurakaré above). But a following full adjacency pair can also be used to resolve problems arising from the SPP, including an other-initiated repair sequence, or in other ways follow up on an SPP.

5.2 Universals of human interaction

The findings of this study suggest that the organization of action-sequencing in human interaction has a remarkable cross-cultural uniformity. Meanwhile, very stable patterns across languages have also been observed in the ways in which turn-taking is organized and the temporal properties it exhibits (Stivers et al. 2009). Research on other-initiated repair has shown very specific parallels in the shape and types of the initiators, the conditions under which different formats are used, and the procedures for resolution of troubles in hearing or understanding (Dingemanse et al. 2015). These three domains form the pillars of conversation analysis and have been identified by Schegloff (2006) as crucial components of the infrastructure for social interaction.

In all three domains, things could have been otherwise, as a quick glance at culturally shaped patterns of institutional exchange reveals. As we noted, we do not expect to find the same kind of sequence organization in courtrooms, presidential press interviews, or classrooms, let alone in hunter-gatherer rituals. Nor do we find the same turn-taking system, nor the same repair system. The universal infrastructure seems to hold specifically in the most informal, ordinary interactive language use, namely conversation. The institutional exchange systems seem to have selected just parts of this broader informal system and specialized it for special cultural purposes (see Drew & Heritage 1992). But the underlying informal system looks culturally independent to a large degree.

This underlying infrastructure for human interaction is part of a raft of special interactional properties that has been called the 'interaction engine' (Levinson 2006, Levinson 2019), which would include the cooperative instincts, the interest in other minds, and the social motivations that propel interaction. Many of these properties are missing or only weakly attested in our nearest cousins, the other great apes (see Section 5.3). It is the interaction engine that affords language, not the other way around – in human ontogeny one observes the interaction engine in proto-conversation in the pre-verbal infant, and it is these abilities that make the acquisition of language possible. We also see the interaction engine at work in the very construction of linguistic systems as in 'home sign' (Goldin-Meadow 2005) or 'cross-signing' when deaf persons from different cultures form a new intermediate language (Byun et al. 2014). Most likely

these abilities also preceded language in phylogeny (Levinson & Holler 2014, Levinson 2016).

The interaction engine not only makes possible our elaborated communications systems, it also makes possible human social systems. Human interaction is shot through with social considerations – we have seen this in the sequence domain with the way in which structures are partly motivated to test sensitive waters, as when a pre-request is used to feel out the probabilities of a request being successful, or when a post-expansion gently queries a disaffiliative SPP, allowing for an ameliorative second version (Schegloff 2007:154). From this we learn that the structures of human social interaction are partly built around the attempt to maintain affiliative, face-preserving behavior (Brown & Levinson 1987), which is itself motivated by the need to maintain the cooperation that makes joint action possible. The discovery of strong universals in sequence organization further strengthens a compelling story about the source of human elite cultural capacities in a universal interactional infrastructure.

5.3 The roots of sequence organization

Both the base adjacency pair and its three positions for expansion sequences, we have argued, can be seen to be motivated by functional considerations. But the existence of other speech exchange systems (as in the various cultural configurations of institutional talk) makes clear that this is not the only conceivable system, and it is worth asking what kinds of origins may lie behind the organization of action-sequencing in conversation.

We noted that the base adjacency pair can be observed in pre-linguistic ontogeny, but interestingly it may also be visible in our nearest great ape cousins, and thus plausibly in phylogeny. Great apes, such as orangutans and bonobos, produce sequences of action remarkably similar to the adjacency pairs found in human interaction (e.g., Rossano 2013, Fröhlich et al. 2016). These sequences consist of two actions, an initiation and a response, much like the first and second pair-parts of adjacency pairs. Indeed, the absence of a response can lead the initiator to repeat the first action, thereby apparently pursuing a response, again much like the pursuits observed in human interaction.

The implication is clear: one of the most basic units of human social interaction, which we have here shown to occur universally across a diverse sample of languages and cultures, has apparent homologs in the social interaction of nonhuman primates. But what about the more elaborate structures that occur in human interaction? Consider, for example, generic pre-expansions. To determine whether chimpanzees would use a first gesture to attract the attention of a partner before they produce a second gesture, Liebal, Call, and Tomasello (2004) examined the complete inventory of gesture sequences used by a group of captive chimpanzees. In a corpus of 75 hours of video, the authors found no evidence of such sequences. Rather than produce preliminary actions to attract attention before subsequent gestures, chimpanzees and other great apes have a simpler solution: they move in front of the non-attending partner before they gesture (Liebal, Call, Tomasello, and Pika 2004). They can also use gestures to draw attention to simultaneous embodied displays, such as a play face or an erection, which in turn foreshadow specific activities (Pika et al. 2005). The chimpanzee analog to a generic pre-expansion does not project the relevance of a subsequent communicative action, contingent upon a response by the recipient to the first. It rather directs the recipient to attend to the source of the signal to find its relevance in already available, non-contingent embodied displays.

What, then, sets human pre-expansions apart, aside from the linguistic resources we use to construct them? A good candidate is hierarchical organization. A preliminary

first pair-part not only makes a response relevant; it also projects the relevance of a subsequent first pair-part. Although hierarchical organization has been observed in complex motoric actions of some great apes (e.g. for food preparation; Byrne & Russon 1998), such capacities have not been clearly adapted to the socio-interactional or linguistic domains, as presumably they have in humans (Pulvermüller 2014). According to Conway and Christiansen (2001), "non-human primates appear to be limited in their ability to learn and represent the hierarchical structure of sequences" (p. 539). Further, the possibility of recursive application of sequence expansion seems well beyond the abilities of other great apes.

5.4 Grammatical adaptations to sequence organization

Understanding the origins and nature of sequence organization is of substantial importance to linguistics. Sequences constitute the contexts in which children learn to use language and discover what utterances achieve. They thus constitute the fundamental bootstrap into language. This phylogenetic and ontogenetic priority must form a substantial set of constraints on linguistic form. Conversation analysts and interactional linguists have argued that the structure of social interaction, as the prime ecological niche of language use, shapes grammar (Schegloff 1996, Ford et al. 2002, Thompson & Couper-Kuhlen 2005, Thompson et al. 2015). The functional pressures of sequence organization, we argue, shape the structure of language in important ways. Consider for example the universality of questions, which have as their central use the function of first parts of adjacency pairs. Indeed, two of the three near universal sentence types (Sadock & Zwicky 1985), interrogatives and imperatives, tend to initiate adjacency pairs. The universality of sequence organization may thus explain the near universality of these forms. Moreover, the precise grammar of interrogation across the world's languages, from intonation, to final particles, to syntactic inversion, may be adapted to the organization of adjacency pairs (see Thompson 1998). The utility of sequence initiation, for example, may explain the cross-linguistically widespread occurrence of tag questions (Ultan 1978), which allow speakers of first pair-parts to pursue responses that have been withheld (see Extract 2) and can retroactively transform actions that do not initiate sequences into ones that do (Sacks et al. 1974:718).

The interactional contingencies of action-sequencing in conversation can also forge pathways for the emergence and grammaticalization of syntactic constructions: left dislocations in conversation often take the form of pre-sequences (Geluykens 1989, 1992), and right dislocations emerge as speakers repair indexical references in pursuit of response (Pekarek Doehler 2011, Bolden et al. 2012). The selection among alternative syntactic constructions for particular actions (e.g., offers) is also sensitive to the structure of sequences, with different positions mandating different forms (Curl 2006). The familiar pattern of noun phrase reduction and anaphora over subsequent mention is in fact closely tied to sequence organization as well: full noun phrases tend to occur in turns that initiate sequences whereas pronouns occur in subsequent turns within the same sequence (Fox 1987). Even the prosody of turns at talk reflects their position within a sequences, with high pitch onsets observed in turns that initiate sequences (Couper-Kuhlen 2004). In all these ways the organization of action sequences in interaction shapes the forms of language that speakers use. Working out the full consequences of the functional pressures of sequence organization on linguistic form should be a prime research objective for linguistic theory.

6   Acknowledgements

7   Appendix

7.1   Sequence annotations symbols

F	First pair-part
S	Second pair-part
F_b	Base first pair-part
S_b	Base second pair-part
F_{pre}	First pair-part of pre-expansion
S_{pre}	Second pair-part of pre-expansion
F_{ins}	First pair-part of insert expansion
S_{ins}	Second pair-part of insert expansion
F_{post}	First pair-part of post-expansion
S_{post}	Second pair-part of post-expansion
SCT	Sequence-closing third
$F_{pursuit}$	First pair-part in pursuit of response

7.2   Transcription conventions

The transcripts employ the following conventions developed by Jefferson (2004).

:	Prolongation of the immediately prior sound
.	Final falling pitch
,	Final intermediate pitch
?	Final rising pitch
[	The point of overlap onset
-	Cut-off of the immediately prior utterance
(0.0)	The duration of gaps and pauses in tenths of seconds
(.)	A brief interval of about a tenth of a second within or between utterances
°	Relatively low volume
=	No break or gap between lines
↑	Shift to especially high pitch
(( ))	Transcriber's descriptions

7.3   Abbreviations for interlinear glosses

1	first person
2	second person
3	third person
A	absolutive
ACC	accusative

ANTIP	antipassive
APPL	applicative
ASSO	assosiative
CVB	converb
COOP	cooperative
COP	copula
DAT	dative
DECL	declarative
DEM	demonstrative
DIM	diminutive
DISTR	distributive
DIT	ditransitive
E	ergative
EMPH	emphatic
EXL	exclamative
FOC	focus
HES	hesitation
ICP	incompletive
IMM	immediate
IMP	imperative
INSTR	instrumental
INTJ	interjection
LOC	locative
N.EGO	non-egophoric
NEG	negative
NMLZ	nominalizer/nominalization
NPST	nonpast
PASS	passive
PRF	prefect
PRT	particle
PFV	perfective
PL	plural
POSS	possessive
PRF	perfect
PROG	progressive
PRS	present
PSN	personal name
PST	past
Q	question particle or affix
QUOT	quotative
SCL	subject clitic
SG	singular
TOP	topic marker

8 Appendix B: Supplementary materials

The supplementary materials for this article can be found online:
https://pure.mpg.de/pubman/item/item_3246830_3/component/file_3246834/Kendrick_etal_2020suppl_Sequence%20organization.pdf

9 References

Atkinson, J. Maxwell, and Paul Drew. 1979. *Order in court: The organisation of verbal interaction in judicial settings*. Atlantic Highlands, NJ: Oxford Socio-Legal Studies.

Bolden, Galina B.; Jenny Mandelbaum; and Sue Wilkinson. 2012. Pursuing a Response by Repairing an Indexical Reference. *Research on Language & Social Interaction* 45.137–155. doi:10.1080/08351813.2012.673380.

Brown, Penelope. 1998. Conversational Structure and Language Acquisition: The Role of Repetition in Tzeltal. *Journal of Linguistic Anthropology* 8.197–221. doi:10.1525/jlin.1998.8.2.197.

Brown, Penelope. 2010. Questions and their responses in Tzeltal. *Journal of Pragmatics* 42.2627–2648. doi:10.1016/j.pragma.2010.04.003.

Brown, Penelope, and Stephen C. Levinson. 1987. *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.

Bruner, Jerome S. 1975. The Ontogenesis of Speech Acts. *Journal of Child Language* 2.1–19. doi:10.1017/S0305000900000866.

Byrne, Richard W., and Anne E. Russon. 1998. Learning by imitation: A hierarchical approach. *Behavioral and Brain Sciences* 21.667–684.

Byun, Kang-Suk; A. Bradford; U. Zeshan; Stephen C. Levinson; and Connie de Vos. 2014. First encounters: Repair sequences in "cross-signing." Poster presented at The International Summer School: Current Issues in Sign Language Linguistics (CISL), Prague.

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

Clark, Herbert H. 1996. *Using language*. Cambridge University Press Cambridge.

Clift, Rebecca. 2016. *Conversation Analysis*. Cambridge: Cambridge University Press.

Conway, Christopher M., and Morten H. Christiansen. 2001. Sequential learning in non-human primates. *Trends in Cognitive Sciences* 5.539–546. doi:10.1016/S1364-6613(00)01800-3.

Coulter, J. (1983). Contingent and A Priori Structures in Sequential Analysis. *Human Studies*, 6(4), 361–376.

Couper-Kuhlen, Elizabeth. 2004. Prosody and sequence organization in English conversation. *Sound patterns in interaction*, ed. by Elizabeth Couper-Kuhlen and Cecilia E. Ford, 335–376. Amsterdam: John Benjamins Publishing Company.

Curl, Traci S. 2006. Offers of assistance: Constraints on syntactic design. *Journal of Pragmatics* 38.1257–1280.

de Vos, Connie; Francisco Torreira; and Stephen C. Levinson. 2015. Turn-timing in signed conversations: coordinating stroke-to-stroke turn boundaries. *Frontiers in Psychology* 268. doi:10.3389/fpsyg.2015.00268.

Dingemanse, Mark. 2015. Other-initiated repair in Siwu. *Open Linguistics* 1, 232-255. doi:10.1515/opli-2015-0001.

Dingemanse, Mark; Seán G. Roberts; Julija Baranova; Joe Blythe; Paul Drew; Simeon Floyd; Rosa S. Gisladottir; et al. 2015. Universal Principles in the Repair of Communication Problems. *PLoS ONE* 10.e0136100. doi:10.1371/journal.pone.0136100.

Drew, Paul, and John Heritage. 1992. Analyzing talk at work: An introduction. *Talk at work: Interaction in institutional settings*, ed. by Paul Drew and John Heritage, 3–65. Cambridge: Cambridge University Press.

Enfield, N. J.; Mark Dingemanse; Julija Baranova; Joe Blythe; Penelope Brown; Tyko Dirksmeyer; Paul Drew; et al. 2013. Huh? What? – A first survey in 21

languages. *Conversational repair and human understanding*, ed. by Makoto Hayashi, Geoffrey Raymond, and Jack Sidnell, 343–380. New York: Cambridge University Press.

Enfield, N. J., Stivers, T., Brown, P., Englert, C., Harjunpää, K., Hayashi, M., Heinemann, T., Hoymann, G., Keisanen, T., Rauniomaa, M., Raymond, C. W., Rossano, F., Yoon, K.-E., Zwitserlood, I., & Levinson, S. C.. Polar answers. *Journal of Linguistics*, 1–28. https://doi.org/10.1017/S0022226718000336

Evans, Nicholas, and Stephen C. Levinson. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32.429–448. doi:10.1017/S0140525X0999094X.

Ford, Cecilia E.; Barbara A. Fox.; and Sandra A. Thompson. 2002. Constituency and the grammar of turn increments. *The language of turn and sequence*, ed. by Cecilia E. Ford, Barbara Fox, and Sandra A. Thompson, 14–38. Oxford: Oxford University Press.

Fox, Barbara. 1987. *Anaphora and the structure of discourse*. Cambridge: Cambridge University Press.

Fröhlich, M.; P. Kuchenbuch; G. Müller; B. Fruth; T. Furuichi; R. Wittig; and Simone Pika. 2016. Unpeeling the layers of language: Bonobos and chimpanzees engage in cooperative turn-taking sequences. *Scientific Reports*, to appear.

Geluykens, Ronald. 1989. The Syntactization of Interactional Processes: Some Typological Evidence. *Belgian Journal of Linguistics* 4.91–103.

Geluykens, Ronald. 1992. *From Discourse Process to Grammatical Construction: On Left-Dislocation in English*. Amsterdam: John Benjamins Publishing Company.

Goffman, Erving. 1976. Replies and Responses. *Forms of Talk*, 5–77. Philadelphia: University of Pennsylvania Press.

Goffman, Erving. 1978. Response Cries. *Language* 54.787–815. doi:10.2307/413235.

Goffman, E. (1981). Footing. In *Forms of Talk* (pp. 124–157). University of Pennsylvania Press.

Goldin-Meadow, Susan. 2005. *The Resilience of Language: What Gesture Creation in Deaf Children Can Tell Us About How All Children Learn Language*. 1 edition. Psychology Press.

Heritage, John. 1984a. A change-of-state token and aspects of its sequential placement. *Structures of Social Action: Studies in Conversation Analysis*, 299–345. Cambridge: Cambridge University Press.

Heritage, John. 1984b. *Garfinkel and Ethnomethodology*. Cambridge: Polity Press.

Hilbrink, Elma E.; Merideth Gattis; and Stephen C. Levinson. 2015. Early developmental changes in the timing of turn-taking: a longitudinal study of mother–infant interaction. *Frontiers in Psychology* 1492. doi:10.3389/fpsyg.2015.01492.

Hockett, Charles F. 1960. The origin of speech. *Scientific American* 203.88-96.

Hoymann, Gertie. 2016. Asking in Freedom: Questioning and requesting in ǂĀkhoe Haiǁom. Nijmegen: Max Planck Institute for Psycholinguistics dissertation, ms.

Hymes, Dell. 1974. *Foundations in Sociolinguistics: An ethnographic approach*. Philadelphia: University of Pennsylvania Press.

Jefferson, Gail. 2004. Glossary of transcript symbols with an introduction. *Conversation analysis: Studies from the first generation*, ed. by Gene H. Lerner, 13–31. Amsterdam: John Benjamins Publishing Company.

Kendrick, K. H., & Torreira, F. (2015). The Timing and Construction of Preference: A Quantitative Study. Discourse Processes, 52(4), 255–289.

Kevoe-Feldman, Heidi, and Jeffrey D. Robinson. 2012. Exploring Essentially Three-Turn Courses of Action: An Institutional Case Study with Implications for Ordinary Talk. *Discourse Studies* 14.217–241. doi:10.1177/1461445612439958.

Lerner, G. H. (2003). Selecting next speaker: The context-sensitive operation of a context-free organization. *Language in Society*, *32*(02), 177–201.

Lerner, Gene. H. (2019). When Someone Other than the Addressed Recipient Speaks Next: Three Kinds of Intervening Action After the Selection of Next Speaker. *Research on Language and Social Interaction*, *52*(4), 388–405.

Levinson, Stephen C. 1983. *Pragmatics*. Cambridge: Cambridge University Press.

Levinson, S. C. (1988). Putting linguistics on a proper footing: Explorations in Goffman's participation framework. In P. Drew & A. Wootton (Eds.), *Goffman: Exploring the interaction order* (pp. 161–227). Polity Press.

Levinson, Stephen C. 2000. *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. Cambridge, MA: MIT Press.

Levinson, Stephen C. 2006. On the human "interaction engine." *Roots of Human Sociality: Culture, Cognition and Interaction*, ed. by N. J. Enfield and Stephen C. Levinson, 39–69. Oxford: Berg.

Levinson, Stephen C. 2013a. Action formation and ascription. *The Handbook of Conversation Analysis*, ed. by Jack Sidnell and Tanya Stivers, 101–130. Malden: Blackwell Publishing Ltd.

Levinson, Stephen C. 2013b. Recursion in Pragmatics. *Language* 89.149–162.

Levinson, Stephen C. 2016. Turn-taking in Human Communication – Origins and Implications for Language Processing. *Trends in Cognitive Sciences* 20.6–14. doi:10.1016/j.tics.2015.10.010.

Levinson, S. C. (2019). Interactional foundations of language: The interaction engine hypothesis. In P. Hagoort (Ed.), Human language: From genes and brain to behavior (pp. 189-200). Cambridge, MA: MIT Press.

Levinson, Stephen C., and Judith Holler. 2014. The origin of human multi-modal communication. *Philosophical Transactions of the Royal Society B: Biological Sciences* 369.20130302. doi:10.1098/rstb.2013.0302.

Liebal, Katja; Josep Call; and Michael Tomasello. 2004. Use of gesture sequences in chimpanzees. *American Journal of Primatology* 64.377–396. doi:10.1002/ajp.20087.

Liebal, Katja; Josep Call; Michael Tomasello; and Simone Pika. 2004. To move or not to move: How apes adjust to the attentional state of others. *Interaction Studies* 5.199–219. doi:10.1075/is.5.2.03lie.

Pekarek Doehler, Simona. 2011. Emergent grammar for all practical purposes: the on-line formatting of left and right dislocations in French conversation. *Constructions: emerging and emergent*, ed. by Peter Auer and Stefan Pfänder, 45–87. Berlin: De Gruyter.

Pika, Simone; Katja Liebal; Josep Call; and Michael Tomasello. 2005. Gestural communication of apes. *Gesture* 5.41–56. doi:10.1075/gest.5.1.05pik.

Pulvermüller, Friedemann. 2014. The syntax of action. *Trends in Cognitive Sciences* 18.219–220. doi:10.1016/j.tics.2014.01.001.

Raymond, Geoffrey. 2003. Grammar and Social Organization: Yes/no Interrogatives and the Structure of Responding. *American Sociological Review* 68.939–967.

Rossano, Federico. 2013. Sequence organization and timing of bonobo mother-infant interactions. *Interaction Studies* 14.160–189. doi:10.1075/is.14.2.02ros.

Sacks, Harvey. 1992a. *Lectures on Conversation*. (Ed.) Gail Jefferson. Vol. 1. 2 vols. Cambridge: Blackwell Publishers.

Sacks, Harvey. 1992b. *Lectures on Conversation*. (Ed.) Gail Jefferson. Vol. 2. 2 vols. Cambridge: Blackwell Publishers.

Sacks, Harvey; Emanuel A. Schegloff; and Gail Jefferson. 1974. A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language* 50.696–735.

Sadock, Jerrold M., and Arnold M. Zwicky. 1985. Speech act distinctions in syntax. *Language typology and syntactic description*, ed. by Timothy Shopen, 155–96. Cambridge: Cambridge University Press.

Schegloff, Emanuel A. 1967. The first five seconds: The order of conversational opening. Berkeley: University of California, Berkeley dissertation.

Schegloff, Emanuel A. 1968. Sequencing in Conversational Openings. *American Anthropologist* 70.1075–1095.

Schegloff, Emanuel A. 1972. Notes on a conversational practice: Formulating place. *Studies in social interaction*, ed. by David N. Sudnow, 75–119. New York: Free Press.

Schegloff, Emanuel A. 1990. On the Organization of Sequences as a Source of "Coherence" in Talk-in-Interaction. *Conversational Organization and its Development*, ed. by Bruce Dorval, 51–77. Norwood, New Jersey: Ablex Publishing Corporation.

Schegloff, E. A. (1996). Turn organization: One direction for inquiry into grammar and interaction. In E. Ochs, E. A. Schegloff, & S. A. Thompson (Eds.), *Interaction and Grammar* (pp. 52–133). Cambridge University Press.

Schegloff, E. A. (1992). Repair After Next Turn: The Last Structurally Provided Defense of Intersubjectivity in Conversation. *American Journal of Sociology*, 97(5), 1295–1345.

Schegloff, Emanuel A. 2006. Interaction: The infrastructure for social institutions, the natural ecological niche for language, and the arena in which culture is situated. *Roots of Human Sociality: Culture, Cognition and Interaction*, ed. by N. J. Enfield and Stephen C. Levinson, 70–96. Oxford: Berg.

Schegloff, Emanuel A. 2007. *Sequence Organization in Interaction: A Primer in Conversation Analysis*. Cambridge: Cambridge University Press.

Schegloff, Emanuel A.; Gail Jefferson; and Harvey Sacks. 1977. The Preference for Self-Correction in the Organization of Repair in Conversation. *Language* 53.361–382.

Schegloff, Emanuel A., and Harvey Sacks. 1973. Opening up Closings. *Semiotica* 8.289–327.

Stivers, Tanya, and Jeffrey D. Robinson. 2006. A preference for progressivity in interaction. *Language in society* 35.367–392.

Stivers, Tanya; and Federico Rossano. 2010. Mobilizing Response. *Research on Language & Social Interaction* 43.3–31. doi:10.1080/08351810903471258.

Stivers, Tanya; N. J Enfield; Penelope Brown; Christina Englert; Makoto Hayashi; Trine Heinemann; Gertie Hoymann; et al. 2009. Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences* 106.10587.

Tomasello, Michael. 2008. *Origins of Human Communication*. Boston: MIT Press.

Thompson, Sandra A. 1998. A discourse explanation for the cross-linguistic difference in the grammar of interrogation and negation. *Case, Typology and Grammar: In honor of Barry J. Blake*, ed. by Anna Siewierska and Jae Jung Song, 309–341. Amsterdam: John Benjamins Publishing Company.

Thompson, Sandra A., and Elizabeth Couper-Kuhlen. 2005. The clause as a locus of grammar and interaction. *Discourse Studies* 7.481.

Thompson, Sandra A.; Barbara A. Fox; and Elizabeth Couper-Kuhlen. 2015. *Grammar in Everyday Talk: Building Responsive Actions*. Cambridge: Cambridge University Press.

Ultan, Russel. 1978. Some general characteristics of interrogative systems.
    Universals of Human Language, Volume 4: Syntax, ed. by Joseph H. Greenberg,
    Charles A. Ferguson, and Edith A. Moravcsik, 211–248. Stanford, CA: Stanford
    University Press.
Walsh, Michael. 1991. Conversational styles and intercultural communication: An
    example from northern Australia. *Australian Journal of Communication* 18.1–12.