Journal of Pragmatics 241 (2025) 1–15

Contents lists available at ScienceDirect

Journal of Pragmatics

journal homepage: www.elsevier.com/locate/pragma

“Please tell me” – The sequential organisation of audience participation in language learning livestreams

Marie Rickert [*], Wyke Stommel

Radboud University, Centre for Language Studies/iHub, the Netherlands

ARTICLE INFO

Article history:
Received 26 September 2024
Received in revised form 24 February 2025
Accepted 26 February 2025
Available online 26 March 2025

Keywords:
Language learning livestreams
Audience participation
Sequential organisation
Self-directed learning
Learning technologies

ABSTRACT

In language learning livestreams, streamers integrate self-directed learning and community engagement with audience members in the chat who may, e.g., assist with language learning activities. This study analyses the sequential organisation of audience participation for language learning support in order to understand joint learning-related practices of streamers and audience members. Our analysis shows that livestreamers particularly recruit learning-related audience assistance when questions arise within their trajectory of self-directed learning. The form of audience participation ranges from brief informings, which frequently occur in larger-size livestreams, to scaffolding sequences in which one audience member engages with the livestreamer over the course of solving a learning task. When moving out of audience participation, livestreamers may perform assessments of the learnable or the learning process and/or praise the audience members who provided help, before moving back to expository talk. Streamers and audience members co-create a new hybrid format between self-directed learning and interactive learning with others who assume positions of peers, teachers, and spectators.

© 2025 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

In diverse societal domains, digital practices emerge and transform social practices. As part of this development, also the language learning landscape has seen a rise in digital learning (Carrier et al., 2017). Digital tools are now an integral part of many onsite classrooms and online courses have gained popularity, especially since the COVID-19 pandemic (Bailey, 2022). Beyond classroom settings, the digital realm also bears affordances for self-directed language learning (Godwin-Jones, 2018). Learners can access a multitude of online language learning resources, make use of learning apps and AI tutors, and participate in online language learning communities. A phenomenon which integrates these different resources are *language learning livestreams*. In such streams, auto-didacts broadcast their self-directed learning activities in real-time for an audience on platforms like YouTube Live or Twitch. Viewers can observe how streamers engage with digital learning materials like videos, e-books, or learning apps as well as with translation tools through screensharing. As is common in livestreams, these activities are accompanied by expository talk; a form of talk in which streamers directly express their associations or demonstrate their skills through chatting along (Tang et al., 2016). In language learning livestreams, expository talk typically involves streamers' comments on the learning activities, which includes evaluative utterances about the language and the learning materials. Expository talk is one way to create watchability: Viewers do not only watch language learning activities, but also listen to streamers’ commentaries that are designed for the audience (Schmidt and Marx, 2020).

---

[*] Corresponding author. P.O. Box 9103, 6500HD Nijmegen, the Netherlands.
E-mail address: marie.rickert@ru.nl (M. Rickert).

https://doi.org/10.1016/j.pragma.2025.02.013
0378-2166/© 2025 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Besides participating as a ratified observer, audience members can also actively contribute to the stream via the chat function, which allows for a form of ‘participatory spectatorship’ (Georgen, 2015). Active involvement of the audience can be initiated in two ways, namely by streamers' recruitment or by audience members' self-selection. The streamers may pick up on contributions from the chat, read them out and engage with them (Song and Licoppe, 2024). In these cases, viewers move from being peripheral participants (‘hearers’, (Goffman, 1979),) to more central participants within the participation framework of the stream. In the scope of language learning streams, they can, for example, provide learning support by explaining vocabulary items, contribute to learning games, or give feedback.

These characteristics make language learning livestreams an innovative format that combines a self-directed approach to language learning with continuous real-time engagement with audience members who may or may not have skills in the target language. Our observations of the live chat reveal that the audience includes native speakers, other learners of the target language, and language-unrelated fans of the livestreamer.

Since livestreaming renders self-directed learning interactive, we are interested in the ways in which the audience provides language learning support in the streamed learning practices. To that end, we employ an ethnomethodological and conversation analytical (EMCA)-perspective to investigate the sequential organisation of audience participation for language learning support (e.g., solving a task, translating an expression, improving pronunciation, etc.). The specific focus is on how such interactions between livestreamers and audience members begin, how language learning support is carried out interactionally by livestreamers and audience members, and how livestreamers situationally close their engagement with the audiences' contributions and move back to self-expository talk. We aim to use the sequential analysis of audience participation as a tool to understand learning-related joint practices of streamers and audience members in language learning livestreams.

2. Self-directed online language learning

Language learning livestreamers are self-directed learners who integrate different forms of online learning and live broadcasting. In general, self-directed online learners are characterised by drawing on a variety of digital resources to pursue their self-study with a high degree of ownership about the studying process and without the guidance of an instructor (Song and Hill, 2007). In contrast to teacher-facilitated co-present learning, self-directed learning requires learners to independently select their learning materials and manage their learning process. This includes seeking clarification—typically provided by teachers in a classroom setting—and maintaining accountability without fixed deadlines or assigned homework.

In the field of self-directed language learning, mobile applications like DuoLingo or Babbel are popular resources, relying on gamification for learning (Sweeney and Moore, 2012). With their drill-like structure and lack of constructive explanatory feedback, such apps have been shown to mainly afford vocabulary acquisition rather than fluency (Heil et al., 2016). Many self-directed language learners furthermore use translation apps like GoogleTranslate or DeepL for vocabulary and pronunciation learning (Van Lieshout and Cardoso, 2022). Such tools can provide translations both in written form and as audio output through text-to-speech technology. A video study of learners' interactions with Google Translate reveals that they employ diverse learning strategies, including listening to the text-to-speech output, repeating the translation, and testing their pronunciation using automatic speech recognition for reverse translation (Van Lieshout and Cardoso, 2022).

Complementing translation apps, many self-directed language learners also picked up on the rise of generative AI, drawing on tools like ChatGPT for language learning. The affordances of such large language models for this purpose include identifying the meaning of a word in context, giving feedback and identifying learners' mistakes, providing translations and examples or even developing quizzes and assignments for learners (Kohnke et al., 2023). In principle, generative AI's aim of creating ‘authentic’, human-like language fits well with language learning tasks (Li et al., 2024). However, learners may also be faced with a ‘knowledge comprehension gap’, i.e., a misalignment between the output of the generative AI and the learners' ability to grasp this output and integrate it with their existing language skills (Li et al., 2023).

While self-directed learners classically carry out their learning trajectory without the guidance of an instructor, some of them engage in online language learning communities, also named ‘polyglot communities’. These communities take the form of social networks, forums or social media groups. There, learners can share resources, discuss learning strategies, set their own learning goals, and hold each other accountable for pursuing these learning goals (Beaven, 2021). In addition, the so-called polyglot community also features content creators whose videos discuss language learning. These content creators classically position themselves as experts who speak multiple languages, demonstrate their skills and share tips on language learning with their followers, whom they address in their videos and who might react in the comment section (Bruzos, 2023). While interactive learning in the classroom targets collaboration through social language practice (Panitz, 1999), these type of online communities treat language as an individual skill rather than as a social practice (Bruzos, 2023).

The format this paper focuses on, language learning livestreams, typically combines different facets of self-directed online learning. To our knowledge, language learning livestreaming has not been the subject of research publications yet. With this study, we hope to contribute insights into this relatively new form of self-directed learning with live community engagement.

3. Interactions between livestreamers and audience members

A small but growing body of literature investigates livestreaming from an EMCA perspective. Livestreaming can cover many different activities such as video-gaming, mukbang (Korean for ‘eating broadcast’), or painting (Recktenwald, 2017).

Interactional research has mainly focused on so-called ‘expository streams’, where livestreamers consistently comment or chat while engaging in activities and audience members contribute via chat (Tang et al., 2016).

One key characteristic of livestreams is their cross-modal character (Jia, 2024; Rosenbaun et al., 2016). Livestreamers engage with audience members through audio and video, while audience members contribute in writing. This cross-modality brings along specific affordances. For instance, a streamer's gaze can indicate their orientation to either the chat or the ongoing activity. If a streamer has not yet looked at the chat and an audience member's first pair part has disappeared due to other incoming messages, the audience member might repost their message once the streamer's gaze shifts towards the chat.

Another interactional feature of livestreams is disrupted turn adjacency, which is also common in computer-mediated communication in general (Herring, 1999). In livestreams, this disruption can occur, for example, during closing sequences (Song and Licoppe, 2023). While livestreams typically feature a four-part-closing-sequence (“general closing, closing responsive message, recipient designed streamer's third-position closing, repetition of generic closing by streamer”, p.13), incoming chat messages with a delay might disrupt this sequential organisation (Song and Licoppe, 2023).

When engaging with the audience, streamers frequently respond to comments that form a first pair part, serving as topic initiations (Licoppe and Morel, 2018). These comments usually consist of greetings or references to local resources observable in the stream (Song and Licoppe, 2024). This type of exchange can be considered ‘upward’ participation mobility (Rosenbaun et al., 2016), as peripheral participants (audience members) address livestreamers and, through their reaction, become central participants. While audience members mostly produce single-turn messages, livestreamers usually respond with multiple, more elaborate turns, often incorporating embodied elements (Recktenwald, 2017). In mukbangs, for instance, audience members might perform text recruitments by soliciting the streamer to eat a specific food item. The streamer may then embody this request by animating the eating action as if they were the requesting viewer, or by ‘puppeteering’—acting “as a puppet, as if moved and controlled by the viewer” (Choe, 2019, p. 21). While these recruitments form the base of a form of co-presence, audience members' recruitments are only acted upon when the livestreamer decides to do so (Choe, 2019).

During a stream, streamers and audience members may co-create and assume different footings (Choe, 2019, 2020). A study on cat mukbangs (eating broadcasts of cats), for example, showed that chat members and livestream hosts may jointly adopt caregiver footings to the cat. However, frequent footing shifts can be observed as audience members, for example, oscillate between caregiver roles, animating the cats, and ‘sport commentator’ roles (e.g., commenting on how a cat is fishing) (Choe, 2020).

While audience-initiated topic initiations are common, streamers may also initiate topics by addressing a first pair part to the audience collectively or to a particular audience member. However, according to Licoppe and Morel (2018), this occurs less frequently. In both cases, the time required for typing a chat contribution and a slight delay in the appearance in the chat due to technological mediation can sometimes disrupt the sequential organisation between the actions and turns of the livestreamer and the chat contributions (Song and Licoppe, 2024). On the other hand, as is characteristic in cross-modal communication, the written mode persists longer which allows livestreamers to read and respond to incoming chat contributions later (Rosenbaun et al., 2016).

The design of audience–streamer interaction is also related to the size of the audience. When the number of viewers remains manageable, livestreamers can engage in conversation-like exchanges with individual viewers. When the audience grows too large, in turn, the chat can take on the form of ‘a roar of a crowd in a stadium’ (Hamilton et al., 2014). In that case, the chat no longer serves as the foundation of a conversation but transforms into what has been described as a ‘waterfall of text’ (Hamilton et al., 2014) or ‘effervescence’ (Song and Licoppe, 2024). Contributions in these ‘waterfalls’ mainly consist of pivoting, i.e., “participant communication with each other that is licensed by the activity and not specifically directed to someone” (Recktenwald, 2017, p. 79). However, these ‘waterfalls’ often contain similar or even identical chat contributions by different audience members, reinforcing each other. As a particular trend emerges in the chat, streamers are more likely to pick up on this trend and, in that sense, ratify the contributions in the ‘waterfall’.

In the context of language learning livestreams, the specific focus of this paper, streamers usually share their screen with learning material, making actions like cursor movement and clicks observable to audience members. Audience members, in their chat contributions, frequently relate to the displayed learning activities, e.g., affirming when a streamer solved a task or providing learning support. Within this cross-modal participation framework, streamers thus move between participatory statuses of ‘speaker’, staging their learning for an audience, and ‘addressed recipient’, receiving messages, e.g., learning support (cf. Dynel, 2014).

4. Methodology

4.1. Motivation

Language learning livestreams present a new form of digital learning. In contrast to many other self-directed learning strategies, they have a cross-modal interactional character. Aiming to understand this new form of digital learning, we use sequential analysis as a tool to access livestreamers' and audience members’ practices, particularly audience participation for language learning support.

4.2. Data collection and analysis

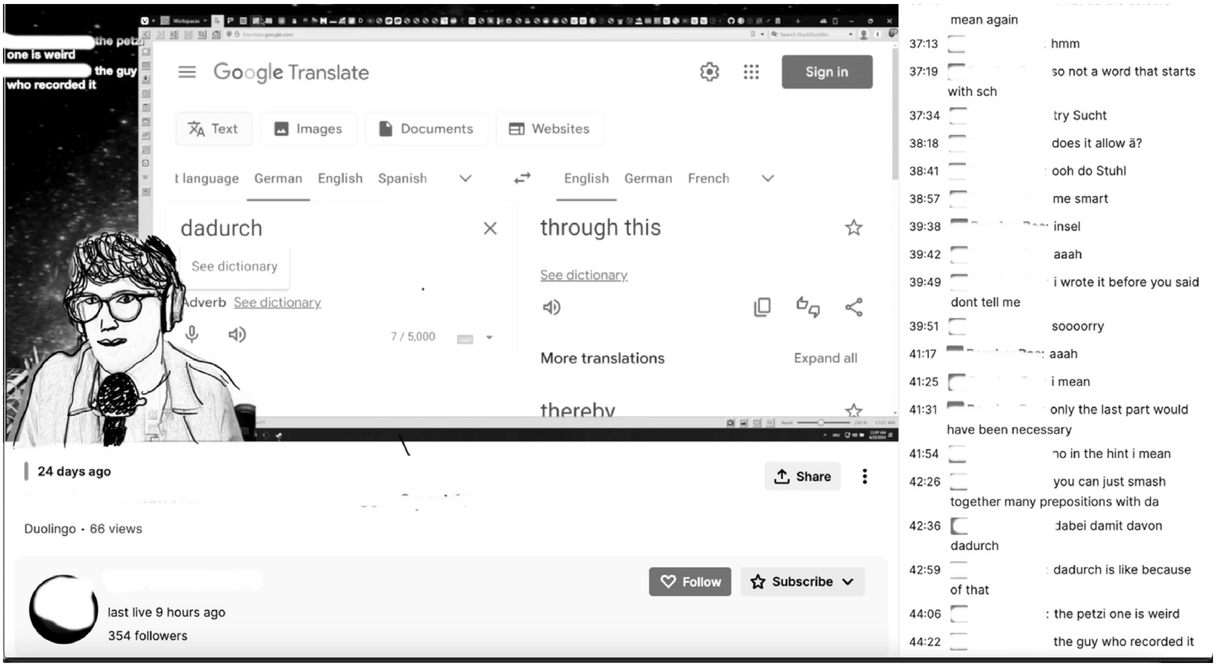
We first conducted an exploratory online search for language learning livestreams for orientation purposes. This search revealed that YouTube Live and Twitch are two relevant platforms for this format. Streamers on Twitch typically do not only focus on language-related content but might also offer, for example, gaming content, while YouTube livestreamers tend to be prominent members of online polyglot communities and also offer content related to language learning other than the livestreams. Also the broadcasted language learning methods differ, with livestreamers on Twitch heavily drawing on language learning apps like DuoLingo and livestreamers on YouTube mainly engaging with learning material like videos or written materials in the target language. Using translation apps or engaging with generative AI is common across the platforms. Since both platforms are popular in their own ways, we decided to include both in this study (two livestreamers per platform). For the analysis, we selected one to two videos per livestreamer based on the following criteria[1]:

(1) Livestreams are held in English.
(2) The target language is a language the first author of this paper (MR) is sufficiently familiar with.
(3) The livestream must have reached more than 50 views (to ensure the relevance of the livestream and chat contributions).

The length of the selected videos varied between 01:17 h and 01:59 h. We carefully watched these videos and compiled an overview of instances in which streamers draw on language learning support from their live audience. At the same time, we made a collection of these instances via the screen-recording function, recording both audio and video, including the chat contributions that appear on the side of the video. This procedure resulted in a collection of 34 cases.

The data (both from YouTube and Twitch) contains the following elements as depicted in Fig. 1:

- The shared screen of the livestreamer, where he[2] often switches between different windows and tabs to deploy different tools of self-directed learning (e.g., videos and books, but also Google Translate and ChatGPT);
- The livestreamer as a “talking head” (Licoppe and Morel, 2012) who can be seen and heard by the audience as he carries out the livestream;
- The parallel chat window in which incoming contributions appear following a timestamp and username (the latter has been deleted in the screenshots for anonymisation purposes). Chat contributions disappear when they are replaced by other chat contributions.


the petzi
one is weird
the guy
who recorded it
Google Translate
Sign in
Text
Images
Documents
Websites
t language
German
English
Spanish
English
German
French
dadurch
See dictionary
Adverb See dictionary
7 / 5,000
through this
See dictionary
More translations
Expand all
thereby
24 days ago
Share
Duolingo • 66 views
Follow
Subscribe
last live 9 hours ago
354 followers

mean again
37:13 hmm
37:19 so not a word that starts with sch
37:34 try Sucht
38:18 does it allow ä?
38:41 ooh do Stuhl
38:57 me smart
39:38 insel
39:42 aaah
39:49 i wrote it before you said dont tell me
39:51 soooorry
41:17 aaah
41:25 i mean
41:31 only the last part would have been necessary
41:54 no in the hint i mean
42:26 you can just smash together many prepositions with da
42:36 dabei damit davon dadurch
42:59 dadurch is like because of that
44:06 : the petzi one is weird
44:22 the guy who recorded it

Fig. 1. Data example from Twitch.

---

[1] In general, one livestream has been selected. In the case of Livestreamer A, two livestreams have been selected since he combines streaming language learning with streaming gaming, leading to a shorter focus on the language learning.

[2] We use the pronoun ‘he’ since all streamers in the data are male.

We analysed the data using an EMCA-approach; inspired by the adaption for video-mediated interaction (Mlynář et al., 2018) and particularly livestreaming (for examples see Choe, 2019, 2020; Jia, 2024; Song and Licoppe, 2023; Song and Licoppe, 2024). In mediated interaction, participants orient to the affordances of the mediating channel (Arminen et al., 2016), requiring us to consider the cross-modal set-up of livestreams when transcribing and analysing. Concretely, this meant that our transcripts and analysis incorporated embodied actions, the livestreamer's shared screen, and incoming chat contributions. We carefully examined the actions preceding and following instances of audience participation recruitment to ensure that no relevant turns were overlooked. This approach was particularly important given that livestreams frequently contain disrupted turn adjacency (Song and Licoppe, 2023).

We transcribed relevant excerpts using a simplified version of Mondada's (2018) transcription conventions which had been adapted for livestreaming (see Licoppe and Morel, 2018 among others). The transcriptions include the streamers' utterances and multimodal actions (e.g., actions visible through screensharing or gaze) as well as incoming chat messages (in bold) in a temporally aligned way. In the transcripts, we analysed the initiation of audience engagement (streamer-initiated or viewer-initiated), how learning support from the audience was situationally carried out, and how the streamer closed the audience engagement sequence. We paid specific attention to the ways in which participants deal with the affordances of the specific media technology when carrying out these actions (e.g., cross-modality between chat and audiovisual streaming with screensharing etc.). As common in CA, we discussed data and shared observations with colleagues during several data sessions.

4.3. Ethical considerations

At the time of data analysis, all livestreams were publicly accessible online, being archived on the streaming platform after they had been broadcasted. Since livestreamers create content for a public audience and chat contributors are presumably aware that their comments appear online, we consider the streams public data and use it for research based on legitimate interest (Art 6.1f GDPR). However, we recognize that both livestreamers and audience members were likely unaware that their contributions to the livestream could be used for research, a key consideration for ethical research involving online data (Stommel and de Rijk, 2021). Therefore, we use pseudonyms and line drawings, reducing recognisability of both the livestreamers and the audience members. Furthermore, we discussed the sensitivity of the data both internally and with members of the Ethical Assessment Board of our faculty. After careful consideration, we concluded that the sensitivity of the data and our focus of analysis are fairly low, so that no direct harm is expected. We received ethical approval for the study by the Ethics Assessment Committee Humanities at Radboud University.

5. Findings

5.1. Recruiting audience assistance

We identified two main types of audience recruitment strategies used by livestreamers to elicit language learning assistance, such as requesting explanations or translations, seeking help with pronunciation, or solving a task. They either directly address specific audience members individually or collectively (e.g., all co-learners in the chat) or make open requests in a general manner to anyone in the chat. In the first type, the audience is usually small so that comments do not appear in a ‘waterfall’-like way (cf. Hamilton et al., 2014). The audience members who are being addressed typically have either demonstrated language skills during the current or a previous livestream, or they have presented themselves as a ‘native speaker’ earlier in the chat.

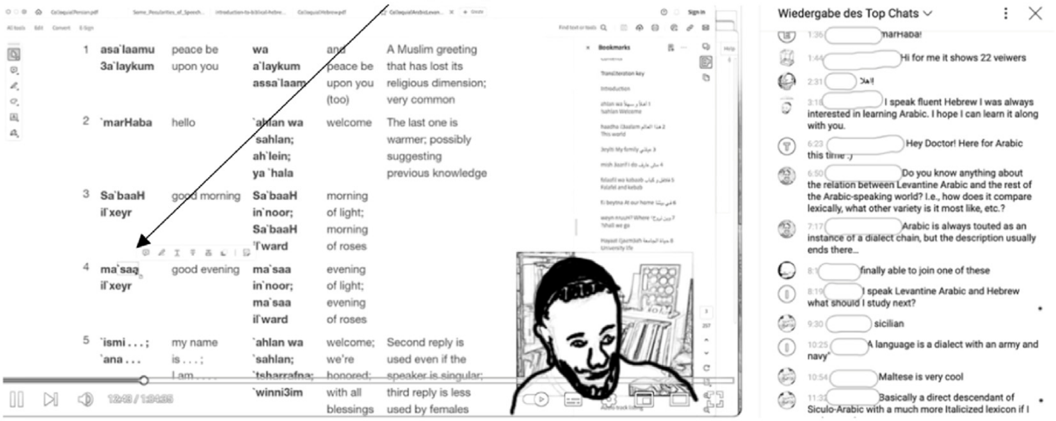
We observed a common orientation to native speakers as capable of giving language support. This orientation becomes visible in an example in which livestreamer 4, a ‘YouTube polyglot’ learning Levantine Arabic, attempts to recruit assistance from native-speaking audience members. He compares the spectrogram of an audio file of the greeting *ma'saa il xeyr*[3] (‘good evening’) with the written transliteration in the textbook, familiarising himself with the pronunciation.

---

[3] The transliteration is taken from the e-textbook that the livestreamer displays.

Exctract 1:

```
01  S     this ehm- accent that they *have here#Fig2
02                                    *highlights “‘saa” in e-textbook
03
```



1	**asa`laamu 3a`laykum**	peace be upon you	**wa a`laykum assa`laam**	and peace be upon you (too)	A Muslim greeting that has lost its religious dimension; very common
2	**`marHaba**	hello	**`ahlan wa `sahlan; ah`lein; ya `hala**	welcome	The last one is warmer; possibly suggesting previous knowledge
3	**Sa`baaH il`xeyr**	good morning	**Sa`baaH in`noor; Sa`baaH 'l`ward**	morning of light; morning of roses	
4	**ma`saa il`xeyr**	good evening	**ma`saa in`noor; ma`saa il`ward**	evening of light; evening of roses	
5	**`ismi . . . ; `ana . . .**	my name is . . . ; I am	**`ahlan wa `sahlan; `tsharrafna; `winni3im**	welcome; we're honored; with all blessings	Second reply is used even if the speaker is singular; third reply is less used by females

Wiedergabe des Top Chats

1:36 [redacted] marHaba!

1:44 [redacted] Hi for me it shows 22 veiwers

2:31 [redacted] اهلا!

3:18 [redacted] I speak fluent Hebrew I was always interested in learning Arabic. I hope I can learn it along with you.

6:23 [redacted] Hey Doctor! Here for Arabic this time :)

6:50 [redacted] Do you know anything about the relation between Levantine Arabic and the rest of the Arabic-speaking world? I.e., how does it compare lexically, what other variety is it most like, etc.?

7:17 [redacted] Arabic is always touted as an instance of a dialect chain, but the description usually ends there...

8:1 [redacted] finally able to join one of these

8:19 [redacted] I speak Levantine Arabic and Hebrew what should I study next?

9:30 [redacted] sicilian

10:25 [redacted] A language is a dialect with an army and navy"

10:54 [redacted] Maltese is very cool

11:32 [redacted] Basically a direct descendant of Siculo-Arabic with a much more Italicized lexicon if I

#Fig2

```
04        eh in the textbook >indicates that< that’s the stressed syllable, so it
05        looks like the stressed syllable. eh: the vowels coalesce (.) and then
06        everything just becomes a syllabic L but *it’s weird to just lose the vowel
07                                                 *clicks to open spectogram
08        in a stressed syllable*
09                                   *clicks to play audio file
10  SPEA  ma’saa il xeyr*
11  S                       *clicks to play audio file
12  SPEA  ma’saa il xeyr*
13  S                       *clicks to stop audio file
14        so if there are any native speakers here eh:m, please let me know if that
15        eh >first of all if< it sounds right and second of all .hhh eh:m you know
16        .hh obviously I think yo- you would >under↑stand it< like ma’saa’il xeyr
17        ((pronounced with glottal stops)) but masaa-il-xeyr ((pronounced tied
18        together)) sounds more natural to me. eh, but I know nothing so eh:
19        >we’ll figure this out< as we go along?
```

In lines 1–8, the livestreamer engages in self-expository talk which consists of him sharing his observations and conclusion about the pronunciation, thereby ‘doing being an expert’ (Kaur, 2011). The livestreamer uses linguistic jargon like ‘the vowels coalesce’ (l.5) and ‘syllabic L’ (l.6), and he multimodally displays the nuanced character of his concern through highlighting a syllable in the e-textbook while claiming that this is ‘the stressed syllable’ (l.1-3). He also delivers ‘proof’ for his observation through opening the spectrogram (l.7) and demonstrates linguistic knowledge through producing a general assessment of the observed phenomenon (“it's weird to just lose the vowel in a stressed syllable”, l.6-8). This multimodal demonstration of being an expert —through both speech and screen activities—is not solely part of the learning activity but is also directed at the audience. As such, it prepares the recruitment of audience assistance from a specific membership category within the otherwise unspecified audience, namely native speakers (“so if there are any native speakers here eh:m, please let me know”, l.14). In this sense, his ‘doing being an expert’ prepares a problem that is linked to his expert opinion, but that he cannot solve on his own. The link is indicated through the “so” (l.14). The ‘if-’construction makes the recruitment conditional, rendering it unproblematic if no answer would appear in the chat. The livestreamer clearly delineates his choice of potentially qualified answerers as he limits his recruitment to native speakers only. He thus seemingly does not see other categories of experts in the audience as suitable to answer his question.

The first part of the address to native speakers concerns the inquiry “if that sounds right” (l.14-15), and the second part involves a juxtaposition of two considered pronunciation options, this time with a personal assessment instead of a general one (l.16-18). This personal assessment ("sounds more natural to me", l.18) is followed by a clear epistemic self-downgrade, namely "but I know nothing" (l.18). In that way, the streamer invites native-speaking audience members to correct him through assigning them a higher epistemic position than himself.

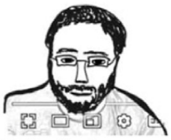
The recruitment fails as he does not receive any comment relating to the pronunciation of *ma'saa' il xeyr* from the chat. As the question has not been answered despite initiation of audience assistance, this allows an orientation to future learning on the side of the livestreamer: The issue is apparently something that neither the livestreamer nor the audience are certain about yet, which the livestreamer orients to with his closing "we'll figure this out as we go along" (l.19).

A more common and less targeted way of seeking help involves addressing all audience members. This is exemplified in the following excerpt, where livestreamer 3, who is learning German, discusses the phrase *Es kommt auf die Situation drauf (sic!) an* ("It depends on it (sic!) the situation"). He encountered this sentence in a German video while streaming. The livestream is very popular with a continuous influx of chat contributions (like a 'waterfall', cf. Hamilton et al., 2014). Presumably linked to the high number of incoming messages, the livestreamer does not attend to the chat while he engages in expository talk just before the fragment starts. He tries to apply the targeted (wrong) grammatical construction to another German sentence in Google Translate while incoming chat messages already point to the incorrectness of the construction.

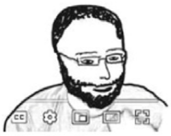
When he remarks that *drauf* ('on it') is missing from the sentence he attempted to translate with Google Translate, he directly addresses the audience, seeking their assistance in determining whether the construction with *drauf* is correct.

Extract 2:

```
01  S    they didn't do it right though. they didn't put the ↑drauf(.)
02       why didn't they put the drauf.
03       could I put the drauf in here guys? >could I say ehm< could
04       could I say <Es kommt auf das Land *DRAUF §an?>*
05                    it depends on it the country
06  uA                                                  §the drauf is wrong
07  S                                              *types drauf between Land and an in
08                                                  GoogleTranslate*
09       #Plea•se        #tell §me
10           •gaze to chat ------------------------------------------------->>
11
```



#Fig3.1



#Fig3.2

```
12  uB                                §Es kommt auf etwas an
13                                     It depends on something
14  uC   ask userX for help a bit
15  uD   same thing
16  uE   Hoe do you learn German the best?
17  uF   drauf is informal
18  uG   drauf is coloquial
19  uH   he drauf is bad
20  S    weiß §es §nicht
21       don't know it
22  uI        §your screen is blurry
23  uJ             §yes it's optional
24  S    the drauf is §wro§ng
25  uK                §Aber du machst das schon ziemlich gut:
26                    §But you already do it quite well:
27  uL                    §drauf is wrong
```

As part of self-expository talk, the livestreamer recognises “they didn't do it right though, they didn't put the *drauf*”, l.1), which forms the base for the ‘noticing-based action’ (Song and Licoppe, 2024) of asking the audience “Could I put the *drauf* in here guys?“, l.3. His insertion of *drauf* into the sentence *Es kommt auf das Land an* (‘It depends on the country’) in Google Translate (lines 7–8) may reflect his self-directed learning strategy, as he frequently relies on Google Translate. At the same time, in the context of audience assistance recruitment, this action reinforces his specific question: *“Could I say ‘Es kommt auf das Land DRAUF an?’ (‘It depends on it the country?’)*".

Simultaneous to the livestreamers’ question, userA already provides an answer in the chat, namely that “the *drauf* is wrong” (l.6). It must be mentioned that user A repeats this brief evaluation here, as they had already pointed out the incorrectness of *drauf* in the chat before the start of the excerpt—at a time when the livestreamer was not yet paying attention to the chat (userA: “The *drauf* is actually wrong there”, data not shown for reasons of space). We found multiple cases in our data where a livestreamer finishes his ongoing project of expository talk before orienting to the chat where audience members had already reacted to earlier events within that stretch of expository talk. Such disrupted turn adjacency is a common feature in livestreams, as it is in computer-mediated interaction more broadly (Herring, 2013). In our data, we observed that livestreamers either pick up on earlier messages from the chat once they have closed their expository talk, or that audience members repeat their chat contributions, sometime several times, until the livestreamer picks up on it. In this case, user A followed the latter strategy, repeating that *“the drauf is wrong”* once the livestreamer addressed the audience with his question.

The livestreamer does not immediately gaze at and read the chat after his question, but first re-enforces his request through the directive “please tell me” (l.9), allowing time for typing and posting chat messages. Shortly after, while he gazes at the chat, more audience members answer his question. These answers largely consist of short evaluations of *drauf* (userF: “drauf is informal”, l.17; userG: “drauf is colloquial”, l.18; userH: “drauf is bad” l.19) and short statements implying that the use of *drauf* is possible but not mandatory (userD: “same thing”, l.15; userJ: “Yes it's optional”, l.23). By means of answering to the livestreamer's question, the audiences' comments inform the livestreamer about the (in-)correctness of *drauf* rather than explaining why the formulation in question is incorrect. The livestreamer makes userA's comment from l.8-9 relevant when reading it out 7 s after its appearance (“the *drauf* is wrong”, l.31). Later (data not shown for reasons of space), he reads out more comments that indicate that *drauf* is colloquial, and eventually summarises “SO HE USED IT but you guys are kind of saying like he used it it's kind of colloquial and stuff”. Reading chat contributions out aloud (Licoppe and Morel, 2018) and summarising them (Lu et al., 2019) is a common practice in livestreams. In this case, summarising links the streamer–audience interaction back to the learning material, indicating that the speaker in the learning video was wrong.

The short form of chat contributions (‘drauf is X’) is common in large audience livestreams, where viewers compete for the streamer's attention (Olejniczak, 2015). In this fast-paced online environment, the streamer often sees individual messages only briefly due to the rapid influx of chat contributions (Olejniczak, 2015).

As evident in this example, in the context of language learning livestreams with large audiences, audience members tend to favour short, informative responses with a *telling* rather than an *explaning* character as second-pair parts to livestreamers’ recruitment of language learning support.

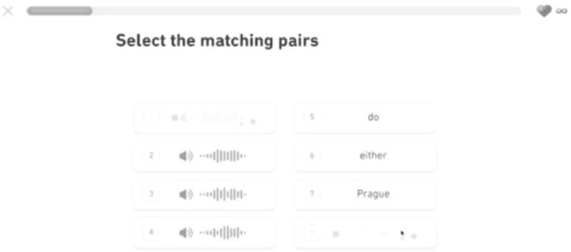
5.2. Audience guidance

As the examples in the previous section demonstrated, the livestreamers themselves frequently direct the course of the livestream: They carry out self-directed learning by drawing on digital learning material like e-textbooks and audio files in the case of livestreamer 4 (Extract 1) or learning videos in the case of livestreamer 3 (Extract 2), and they contextualise their engagement with these resources through expository talk. The previous section demonstrated that the livestreamers may also address the audience directly to recruit their assistance. While the livestreamers are thus mainly in charge of the ongoing activities, there are also instances where the audience takes a more active role in shaping the course of activities and, consequently, the learning process.

This is the case in Extract 3. German-learning livestreamer 1 plays DuoLingo and, in the scope of one exercise, matched the German vocabulary item ‘*fließend*’ with its English translation ‘fluent’. Thirty nine seconds later, while the streamer has proceeded to other tasks, an audience member links back to the earlier task posting ‘*Fließende Flüsse* [flowing rivers, with ‘*fließend*’ translating to ‘flowing’ rather than ‘fluent‘ in this case]’. The livestreamer reads out the comment and interrupts his DuoLingo game to investigate the meaning of the chat contribution:

Extract 3:

```
01 uM   Fließende Flüsse*
   S                   *clicks on audio 1
02 SPEA ändern
03 S    ändern is chang*ing #
04                    *clicks on option ‘change’ in DuoLingo
05
```



```
                             #Fig4
06 S    *clicks on audio 2
07 SPEA un•ternehmen
08 S      •-gaze to chat---•
09                   unter->•nehmen is doing* un•dertaking?
10                           •---gaze to screen-•
11                                   *clicks on option ‘do’ in DuoLingo
12                                      •----gaze to chat--------------->
13      I like to think of it as (.8) fließende (.) Flüsse. fluent (.) rivers?•>
14      (1.8) know what that is? (1.9) fließende (.) Flüsse (1.9)*
15                                                        *switches tab to
16                                                         GoogleTranslate
17      *types ‘fließende Fl§ü§sse’, translation ‘flowing rivers’ appears
18                            §let’s see
19 uM                            §just wanted to point out that fließen does not only
                                  mean fluent with that
20 S    flowing river*s
21                   *clicks speech-to-text-function on GoogleTranslate
22 SPEA fließende Flüsse
23 S    WHAT? (2.1) oh goo:d! oh wow, I had no idea?
```

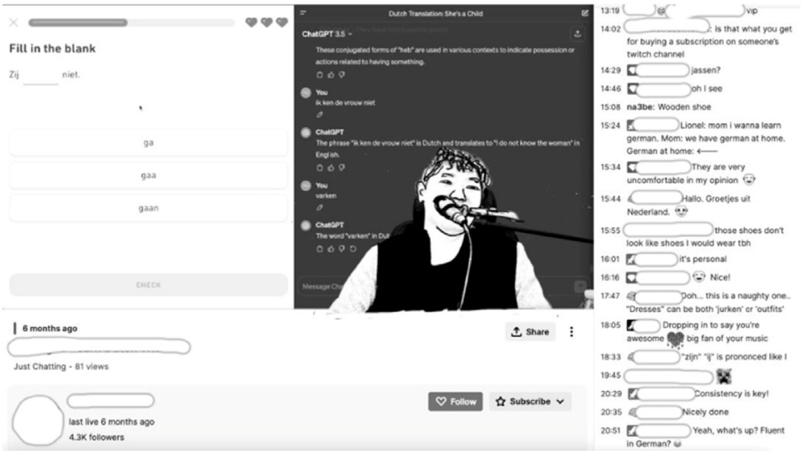
A frequently occurring format for livestreamers to process incoming chat contributions is the *‘read aloud and respond’*-practice (RAR) (Licoppe and Morel, 2018). In the case at hand, the livestreamer reads aloud userL's incoming chat contribution ‘*fließende Flüsse’* (l.13) and responds by demonstrating his non-understanding (“know what that is?”, l.14). The response-unit of the sequence is extended as the livestreamer then starts looking up the meaning of userL's contribution on Google Translate (l.15-17). While the livestreamer does so, userL posts an account for bringing up the expression ‘*Fließende Flüsse*’ in the first place (“just wanted to point out that *fließen* does not only mean fluent”, l.19). The result on Google Translate indeed provides the translation ‘flowing’ (l.20), which the livestreamer reads out and then listens to, signalling serious engagement with userL's input. Finally, the news receipt markers “what?“, and “oh” prefacing the positive assessments “good” and “wow” (l.23), further contribute to framing userL's contribution as valuable. In this case, ‘oh’ functions as a change-of-stake-token used to accept the prior message as informative (Goodwin and Heritage, 1990). Moreover, the addition “I had no idea?” (l.23) constructs the audience member's contribution as having led to a learning moment.

One form of audience guidance thus concerns the suggestion of learnables from the audience. While the explanation of the learnable by the audience member was minimal in Extract 3, there are also instances where audience members offer more detailed explanations, temporarily assuming a teacher-like position in the learning process.

This is the case in the following extract, taken from a medium-sized livestream with under 15 active chat participants. Dutch-learning livestreamer 2 is trying to solve a DuoLingo task requiring him to conjugate the verb ‘*gaan* [to go]’ in the third person plural. The task shows the sentence ‘*Zij* __________ *niet* [They do not ________]’ and presents three different options to fill in the blank ([ ] *ga*; [ ] *gaa*; [ ] *gaan*). The livestreamer first draws on ChatGPT to help him solve the task:

Extract 4:

```
01
```



Zij ______ niet.

These conjugated forms of "heb" are used in various contexts to indicate possession or actions related to having something.

You
ik ken de vrouw niet

ChatGPT
The phrase "ik ken de vrouw niet" is Dutch and translates to "I do not know the woman" in English.

You
varken

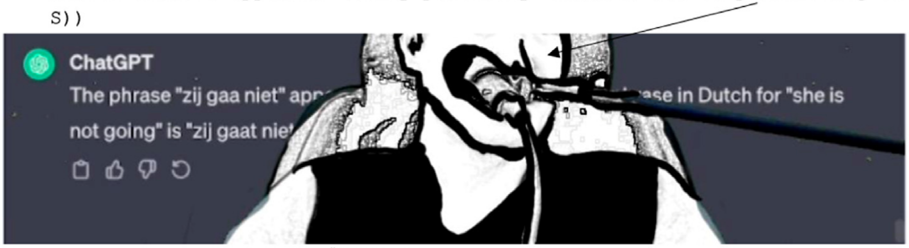
ChatGPT
The word "varken" in Du

6 months ago
[redacted]
Just Chatting · 81 views

[redacted]
last live 6 months ago
4.3K followers

13:19 [redacted] @[redacted] vip
14:02 [redacted]: is that what you get for buying a subscription on someone's twitch channel
14:29 [redacted] jassen?
14:46 [redacted] oh I see
15:08 na3be: Wooden shoe
15:24 [redacted] Lionel: mom i wanna learn german. Mom: we have german at home. German at home: ⟵
15:34 [redacted] They are very uncomfortable in my opinion
15:44 [redacted] Hallo. Groetjes uit Nederland.
15:55 [redacted] those shoes don't look like shoes I would wear tbh
16:01 [redacted] it's personal
16:16 [redacted] Nice!
17:47 [redacted] Ooh... this is a naughty one.. "Dresses" can be both 'jurken' or 'outfits'
18:05 [redacted] Dropping in to say you're awesome big fan of your music
18:33 [redacted] "zijn" "ij" is prononced like I
19:45 [redacted]
20:29 [redacted] Consistency is key!
20:35 [redacted] Nicely done
20:51 [redacted] Yeah, what's up? Fluent in German?

```
        Fig5.1
02      *types zij §gaa niet in ChatGPT------*
03  S             §I'm gonna ask§(0.5)§(1.4)*
04  uK                           §Zij is a mean one as well
05  uL                                 §Lol
06      if +this is right
07         +answer ChatGPT appears# ((only partially readable due to positioning of
08          S))
09
```



ChatGPT
The phrase "zij gaa niet" app ase in Dutch for "she is
not going" is "zij gaat niet

```
                           #Fig5.2
10      gaat (2.4)§ (1.9)
11  uK            §can be "she" OR "they"
12  S   she is not going (2.8) what? The ga is not even •on this. Okay, help• Ehehe:
13                                                      •---gaze to screen--•
14      It's gotta be either §gaat•-----(3)----•
15  uK                      §"Gaa" is not a thing lol
16                              •gaze to chat•
17  S   okay e.hh. Is *-------it ga------* *------or gaan?------*
18                    *moves cursor to ga* *moves cursor to gaan*
19      I think it's gaan right userK? 'cause ChatGPT is not §helping me
20  uK                                                        §She is plural in this
                                                               case
21  S   .hhhh•---(2.4)----•(.)
22           •gaze to chat•

23  uK  so the plural of 'to go'
24  S   Okay! (.)•It could be-• they.•so it's gaan!(2.9) right G-A-A-•N. Gaan?(2)§ •
25               •gaze to chat•      •------------gaze to chat---------------------•
26  uK                                                                            §yup
27  S   Ehe: ok*ay! nice nice nice! eh+e::
28             *clicks on gaan
29                                 +DuoLingo 'ping' sounds and gaan lights up
30                                  green
```

In the beginning of this sequence, the streamer displays that he has problems with the task at hand as he prompts ChatGPT to indicate if *gaa* is the correct option (l.2). On a continuum of recruitments (Kendrick and Drew, 2016), such a display could be seen as a softer form of recruitment as it leads user K to anticipate the livestreamer's need for assistance. He evaluates the task ("*Zij* is a mean one as well", l.4) without it being picked up by the livestreamer. ChatGPT fails to provide the correct answer in the context of the task since it bases the conjugation on *zij* ('she/they') as third person singular instead of third person plural. On the side of the livestreamer, trouble manifests due to a knowledge comprehension gap (Li et al., 2024) when he reads out loud the conjugation by ChatGPT (l.10), followed by a 4.1 s silence and the reading out loud of a part of ChatGPT's explanation ("she is not going", l.12). While the silence already indicates non-understanding, this is confirmed when he notices ("what? the *gaa* is not even on this!", l.12) and then proceeds to shift his gaze to the chat (l.13), followed by a recruitment of audience support through the directive "okay help!" (l.12).

In the chat, userK already indicated "can be 'she' or 'they'" (l.11) after the livestreamer demonstrated difficulties with the task to the audience through prompting ChatGPT. In this medium-sized livestream, contrary to the common practice in large audience language learning livestreams (see Extract 2), the user does thus not immediately provide the correct answer to the livestreamer. Instead, he initiates a scaffolding sequence. Scaffolding is an interactional phenomenon in teaching and learning, which involves temporary support to solve a task, tailored to the learners' level, which in turn is situationally expressed by the learner (Koole and Elbers, 2014; van de Pol et al., 2010). In this context, his next chat contribution "*Gaa* is not a thing lol" (l.15) can be seen as a second pair part to the livestreamers' prior wondering about *gaa* (l. 12). The livestreamer is responsive to userK's scaffolding element and eliminates *gaa* from the options on his screen ("is it *ga* or *gaan*" l.17). His question for confirmation of his answer of choice to userK (l.19) contrasts his earlier general recruitment of help that was not directed to anyone in specific, showing that he now treats userK as an epistemic authority on this task. Instead of confirming the livestreamers' answer, userK provides two more hints ("she is plural in this case", l.20; "so the plural of to go", 23). UserK only confirms the livestreamer's repeated question for confirmation once the livestreamer demonstrated an understanding of the correct answer through relating to his hints ("It could be they", l.24) and drawing the correct conclusion ("so it's gaan", l.24). Finally, the livestreamer logs his answer on the screen and positively assesses his presumed success (l.27-28) even before DuoLingo confirms the correctness of the answer (l.29-30). Despite the livestreamer's initial first pair parts that preferred a simple confirmation (l.19), the audience member employed scaffolding techniques, guiding the livestreamer in completing the task.

Both in Extract 3 and 4, audience members guide the streamers' learning situationally. However, the initiation of this guidance differs: Whereas an audience member directs the livestreamers' attention to a learnable while the streamer already continued with other learning activities in Extract 3, Extract 4 demonstrates how an audience member offers help when the streamer displays problems with a current task. In both cases, the streamer picks up on the help and incorporates the audience member's input in the course of the learning activity.

5.3. Moving out of audience participation

Once the livestreamer has solved the task at hand or learnt a learnable through audience support, the livestreamer and audience members usually move away from language learning support by reproducing the configuration in which the livestreamer engages in a language learning activity with accompanying expository talk. A common element of this moving out of audience participation are assessments of either the learning moment (as was the case in Extract 3: "oh good", l.24), or the joint solving of the task at hand (as is the case in Extract 4 "nice, nice, nice", l.28).

Audience members may also be praised for their contribution by the livestreamer. In the next extract, German-learning Livestreamer 1 plays German wordle, a popular word puzzle game in which players need to guess a five-letter word. After seven failed attempts by the livestreamer to complete the word 'KRA_ _' with two letters, userL suggests 'Krach (loud noise)' in the chat:

Extract 5:

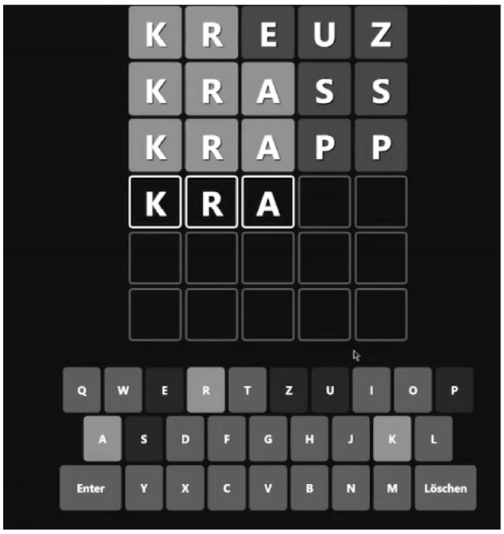


Fig. 6

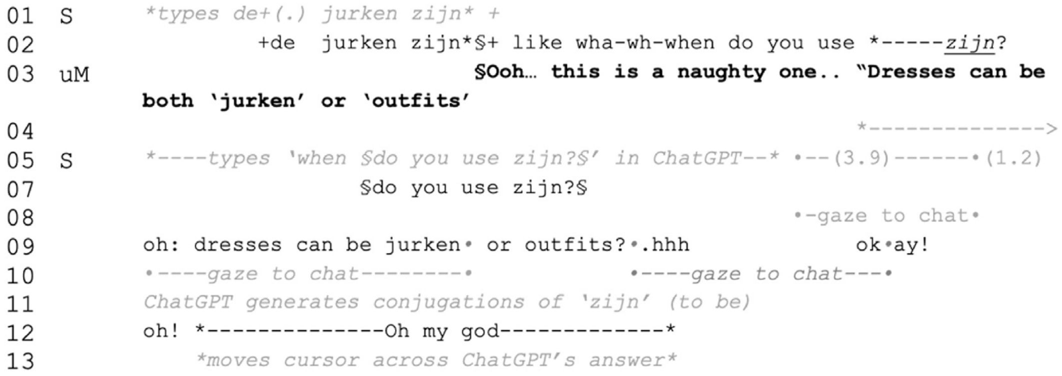
```
01  uL     Krach?*
02  S            *enters KRACL
03                            *removes L, adds K+
04                                              +DuoLingo shakes word (=wrong answer)
05  S      •*removes C-K*•
06         •gaze to chat-•*
07                        *adds C-H*+
08                                +DuoLingo letters turn orange (=right answer)+
09                                +---------K - R - A §- C -H-----------------+§
10  S                                                   §----There we go:oh!-----§
11         userL got it again!*
12                            *clicks to play sound
13  SPEA   Grea§t job! Thank you!
14  uL         §((clapping chick emoji))
15  S      Das war Krach!
16         That was Krach!
```

The livestreamer constructs the success as a joint event in which he himself, userL, and potentially other viewers take part ("There we go", l.10). UserL is praised for his contribution which led to the success ("userL got it again", l.11). This praise consists of the livestreamers' direct mentioning of his name with the emphasis on his repeated success. In combination with the audio file the livestreamer plays, which reads out 'Great job, thank you!' in a distorted voice (l.13), userL's contribution that facilitated the successful completion of the word puzzle is treated as part of the game by the livestreamer. The gamification of learning inherent in wordle thus extends to the interaction between the livestreamer and the audience. Next, the streamer formulates the game solution "Das war *Krach*! [*That was Krach!"* (l.15-16), which has a moderating function.

Ultimately, moving out of audience participation usually entails the livestreamers' observable re-orientation to a means of self-directed learning. The livestreamers, for example, click to continue a video or the next task appears on screen, which they then engage with. The following extract shows how an event on the screen, which is also visible to the audience due to screensharing, affords the livestreamers' orientation away from the audience without requiring further explanation.

Dutch-learning Livestreamer 3 is prompted to translate the sentence "The dresses are good" by DuoLingo. He expresses his unknowingness and types the sentence into an open thread 'Dutch translations' in ChatGPT. After receiving ChatGPT's translation, he proceeds to enter it into Duolingo.

Extract 6:



```
01 S     *types de+(.) jurken zijn* +
02               +de  jurken zijn*§+ like wha-wh-when do you use *-----zijn?
03 uM                               §Ooh… this is a naughty one.. "Dresses can be
         both 'jurken' or 'outfits'
04                                                                     *-------------->
05 S     *----types 'when §do you use zijn?§' in ChatGPT--* •--(3.9)------•(1.2)
07                         §do you use zijn?§
08                                                                •-gaze to chat•
09       oh: dresses can be jurken• or outfits?•.hhh                     ok•ay!
10       •----gaze to chat--------•            •----gaze to chat---•
11       ChatGPT generates conjugations of 'zijn' (to be)
12       oh! *---------------Oh my god-------------*
13          *moves cursor across ChatGPT's answer*
```

While ChatGPT's generates a longer answer, the livestreamer gazes at the chat and notices a contribution by userK. userK's warning, relevant to the task, is acknowledged by the livestreamer through a 'read aloud and respond'-sequence (Licoppe and Morel, 2018) This sequence is prefaced by the news-receipt marker 'oh' ("oh, dresses can be jurken or outfits?.hhh okay", l.9), indicating that this is new information to the livestreamer. The acknowledgment of the information with "okay" closes the audience engagement sequence. Hence, the livestreamer does not evaluate the audience member's contribution (cf. Extract 3), thank userK for the warning (cf. Extract 5) or elaborate, but instead transitions back to expository talk immediately, marked by gazing to ChatGPT and moving his cursor (l.13). Hence, the livestreamer prioritises progressivity of the learning activity he is currently engaged in over further interaction with an audience member.

6. Discussion and conclusion

In this paper, we investigated the sequential organisation of learning support via chat by the audience in language learning livestreams on YouTube and Twitch. We have shown that livestreamers particularly recruit learning-related audience assistance when questions arise within their trajectory of self-directed learning. Audience members may also initiate language learning support, e.g., suggesting a learnable or an answer in the context of learning games. While the audience's language learning support in large audience streams typically consists of informings, audience members in smaller audience streams may engage in scaffolding, which comes closer to assuming a 'teacher' position. When moving out of audience participation, livestreamers may perform assessments of the learnable or the learning process and/or praise the audience member who provided help. Ultimately, moving out of audience participation entails transitioning back to expository talk and orienting to the means of self-directed learning at hand, which audience members may also visually access through the screensharing function.

Via screensharing, the streamers demonstrate their strategies of self-directed learning. Concretely, they combine a variety of means of self-directed learning, including e-books, videos, learning apps like DuoLingo as well as large language models like ChatGPT and translation apps like GoogleTranslate. Learning support by the audience can, for example, be made relevant in case these means fail to deliver an understandable explanation or solution to the task at hand to the livestreamer. This is for example the case when a knowledge comprehension gap emerges (Extract 4: "I think it's *gaan*, right, user K? ChatGPT is not helping me"; also cf. Extract 2). Knowledge comprehension gaps have been described as a limitation of self-directed learning, particularly in the context of generative AI (Li et al., 2024). Thus, the specific media communication of language learning livestreams affords to move from self-directed learning to audience support when problems surface. Livestreamers then recruit audience members for learning support, either through addressing individual audience members or membership categories like native speakers, or through open requests to all viewers.

On the other hand, streamers may draw on means of self-directed learning to help with some aspects of the learning support offered by the audience. For example, they look up a translation of a learnable suggested in the chat and prompt GoogleTranslate's text-to-speech output (Extract 3). Text-to-speech is a commonly used tool to check pronunciation, as had been found by van Lieshout and Cardoso (2022). In the context of learning support during livestreams (without call-in functions), it offers streamers a facet of support (speech) that audience members cannot provide.

All in all, language learning livestreamers use the affordances of different technologies – the chat and different online learning tools - to enhance self-directed learning. Through their interactions with audience members in the chat, they co-create a hybrid format between self-directed learning and interactive learning with others who assume positions of peers, teachers, and spectators. Therefore, language learning livestreams afford new learning interactions characterised through exemplifying demonstrations of learning and entertainment.

Transcription conventions

Based on Licoppe and Morel (2018); Song and Licoppe (2023, 2024): The conventions integrate Mondada's (2018) conventions for embodied action and Jefferson's (2008) conventions for talk.

S: streamer; uX: userX SPEA: speaker
Bold: chat messages
↑: rising intonation
(x.x): timed pause in absolute seconds
::: stretched sound
.hh: outbreath
WORD: loud
– embodied actions, synchronised with corresponding stretches of talk
§: moment when the audiences' message appears
#: moment when screenshot appears
—>: action continues across subsequent line(s)
—>>: action continues after the extracts' end

CRediT authorship contribution statement

Marie Rickert: Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Wyke Stommel:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Funding sources

This work was supported by Ammodo - Foundation for Arts, Architecture and Science.

Declaration of competing interest

The authors declare that there were no financial interests/personal relationships which may be considered potential competing interests.

Data availability

Data will be made available on request.

References

Arminen, I., Licoppe, C., Spagnolli, A., 2016. Respecifying mediated interaction. Res. Lang. Soc. Interact. 49 (4), 290–309.

Bailey, D., 2022. Interactivity during Covid-19: mediation of learner interactions on social presence and expected learning outcome within videoconference EFL courses. J. Comp. Educ. 9 (2), 291–313.

Beaven, A., 2021. Learning without a teacher: self-Directed Language Learning in the digital wilds. In: Beaven, T., Rosell-Aguillar, F. (Eds.), Innovative Language Pedagogy Report. Research-publishing.net.

Bruzos, A., 2023. 'Language hackers': YouTube polyglots as representative figures of language learning in late capitalism. Int. J. Biling. Educ. Biling. 26 (10), 1210–1227.

Carrier, M., Damerow, R.M., Bailey, K.M., 2017. Digital Language Learning and Taching. New York, NY: Routledge.. Cazden, CB (1996). A Pedagogy of Multiliteracies: Designing Social Futures. Harvard Educational Review, 66, pp. 60–92 (1).

Choe, H., 2019. Eating together multimodally: collaborative eating in mukbang, a Korean livestream of eating. Lang. Soc. 48 (2), 171–208. https://doi.org/10.1017/S0047404518001355.

Choe, H., 2020. Talking the cat: footing lamination in a Korean livestream of cats mukbang. J. Pragmat. 160, 60–79. https://doi.org/10.1016/j.pragma.2020.02.009.

Dynel, M., 2014. Participation framework underlying YouTube interaction. J. Pragmat. 73, 37–52. https://doi.org/10.1016/j.pragma.2014.04.001.

Georgen, C., 2015. Well Played & Well Watched: Dota 2, Spectatorship, and eSports. Well Played A Journal on Video Games, Values, and Meaning, 4, pp. 179–191 (1).

Godwin-Jones, R., 2018. Chasing the butterfly effect: informal language learning online as a complex system. Lang. Learn. Technol. 22 (2), 8–27 doi.org/10125/44643.

Goffman, E., 1979. Footing. Semiotica 25 (1–2), 1–30. https://doi.org/10.1515/semi.1979.25.1-2.1.

Goodwin, C., Heritage, J., 1990. Conversation analysis. Annu. Rev. Anthropol. 19 (1), 283–307.

Hamilton, W.A., Garretson, O., Kerne, A., 2014. Streaming on twitch: fostering participatory communities of play within live mixed media. In: CHI '14: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1315–1324. https://doi.org/10.1145/2556288.2557048.

Heil, C.R., Wu, J.S., Lee, J.J., Schmidt, T., 2016. A review of mobile language learning applications: Trends, challenges, and opportunities. EuroCALL Rev. 24 (2), 32–50.

Herring, S.C., 1999. Interactional coherence in CMC. In: Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences. 1999. HICSS-32. Abstracts and CD-ROM of Full Papers.

Herring, S.C., 2013. 10. Relevance in computer-mediated conversation. Pragm. Comp. Mediat. Commun. 9, 245.

Jefferson, G., 2008. Glossary of transcript symbols with an introduction. In: Conversation Analysis: Studies from the First Generation. John Benjamins Publishing Company, pp. 13–31.
Jia, M., 2024. Cross-modal Management of Trolling During Live Streaming on Periscope: A Micro-analysis. Explorations in Internet Pragmatics: Intentionality, Identity, and Interpersonal Interaction, pp. 225–245.
Kaur, J., 2011. Doing being a language expert: the case of the ELF speaker. Latest Trends ELF Res. 5375.
Kendrick, K.H., Drew, P., 2016. Recruitment: offers, requests, and the organization of assistance in interaction. Res. Lang. Soc. Interact. 49 (1), 1–19. https://doi.org/10.1080/08351813.2016.1126436.
Kohnke, L., Moorhouse, B.L., Zou, D., 2023. ChatGPT for language teaching and learning. RELC J. 54 (2), 537–550.
Koole, T., Elbers, E., 2014. Responsiveness in teacher explanations: a conversation analytical perspective on scaffolding. Ling. Educ. 26, 57–69. https://doi.org/10.1016/j.linged.2014.02.001.
Li, B., Bonk, C.J., Kou, X., 2023. Exploring the multilingual applications of ChatGPT: uncovering language learning affordances in YouTuber videos. Int. J. Comput. Assist. Lang. Learn. Teach. 13 (1), 1–22.
Li, B., Bonk, C.J., Wang, C., Kou, X., 2024. Reconceptualizing self-directed learning in the era of generative AI: an exploratory analysis of language learning. IEEE Trans. Learn. Technol. 17, 1489–1503.
Licoppe, C., Morel, J., 2012. Video-in-interaction:“Talking heads” and the multimodal organization of mobile and Skype video calls. Res. Lang. Soc. Interact. 45 (4), 399–429.
Licoppe, C., Morel, J., 2018. Visuality, text and talk, and the systematic organization of interaction in Periscope live video streams. Discourse Stud. 20 (5), 637–665. https://doi.org/10.1177/1461445618760606.
Lu, Z., Annett, M., Wigdor, D., 2019. Vicariously experiencing it all without going outside: a study of outdoor livestreaming in China. Proc. ACM Human Comp. Inter. 3 (CSCW), 1–28.
Mlynář, J., González-Martínez, E., Lalanne, D., 2018. Situated organization of video-mediated interaction: a review of ethnomethodological and conversation analytic studies. Interact. Comput. 30 (2), 73–84.
Mondada, L., 2018. Multiple temporalities of language and body in interaction: challenges for transcribing multimodality. Res. Lang. Soc. Interact. 51 (1), 85–106.
Olejniczak, J., 2015. A linguistic study of language variety used on twitch. tv: desriptive and corpus-based approaches. Redefin. Commun. Intercult. Context 4 (1), 329–334.
Panitz, T., 1999. Collaborative versus Cooperative Learning: A Comparison of the Two Concepts Which Will Help Us Understand the Underlying Nature of Interactive Learning US. Department of Education, Office of Educational Research and Improvement, pp. 3–13. https://files.eric.ed.gov/fulltext/ED448443.pdf.
Recktenwald, D., 2017. Toward a transcription and analysis of live streaming on Twitch. J. Pragmat. 115, 68–81. https://doi.org/10.1016/j.pragma.2017.01.013.
Rosenbaun, L., Rafaeli, S., Kurzon, D., 2016. Participation frameworks in multiparty video chats cross-modal exchanges in public Google Hangouts. J. Pragmat. 94, 29–46.
Schmidt, A., Marx, K., 2020. Making Let's Plays Watchable: An Interactional Approach to Gaming Visualizations. Visualizing Digital Discourse: Interactional, Institutional and Ideological Perspectives, p. 131.
Song, L., Hill, J.R., 2007. A conceptual model for understanding self-directed learning in online environments. J. Interact. Online Learn. 6 (1), 27–42.
Song, L., Licoppe, C., 2023. Closing live video streams: a sequential analysis. Discourse. Context Media 53, 100698.
Song, L., Licoppe, C., 2024. Noticing-based actions and the pragmatics of attention in expository live streams. Noticing ‘effervescence’and noticing-based sequences. J. Pragmat. 226, 1–16.
Stommel, W., de Rijk, L., 2021. Ethical approval: none sought. How discourse analysts report ethical issues around publicly available online data. Res. Ethics 17 (3), 275–297.
Sweeney, P., Moore, C., 2012. Mobile apps for learning vocabulary: categories, evaluation and design criteria for teachers and developers. Int. J. Comput. Assist. Lang. Learn. Teach. 2 (4), 1–16.
Tang, J.C., Venolia, G., Inkpen, K.M., 2016. Meerkat and periscope: I stream, you stream, apps stream for live streams. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems.
van de Pol, J., Volman, M., Beishuizen, J., 2010. Scaffolding in teacher–student interaction: a decade of research. Educ. Psychol. Rev. 22 (3), 271–296. https://doi.org/10.1007/s10648-010-9127-6.
Van Lieshout, C., Cardoso, W., 2022. Google Translate as a tool for self-directed language learning. Lang. Learn. Technol. 26 (1).

Dr. Marie Rickert is a post-doctoral researcher at Radboud University, Centre for Language Studies/iHub (Interdisciplinary Research Hub on Digitalization and Society). Her research focuses on interaction in education, with a specific focus on linguistically diverse contexts.

Prof. Wyke Stommel is professor of Language and Interaction in Digital Contexts at Radboud University, Centre for Language Studies/iHub (Interdisciplinary Research Hub on Digitalization and Society). Her research concerns spoken and digital communication, mainly in institutional environments.