Multimodal noun phrases

Leelo Keevallik

The self-archived postprint version of this journal article is available at Linköping University Institutional Repository (DiVA):

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-168937>

N.B.: When citing this work, cite the original publication.

Keevallik, L., (2020), Multimodal noun phrases, *The 'Noun Phrase' across Languages*, , 154-177.
<https://doi.org/10.1075/tsl.128.07kee>

Original publication available at:

<https://doi.org/10.1075/tsl.128.07kee>

Copyright: John Benjamins Publishing Company

<https://benjamins.com/content/home>

Multimodal “noun phrases”¹

Leelo Keevallik

Linköping university

Abstract

In co-present interaction, out bodies are continuously available for sense-making. Linguists, however, have generally analyzed grammatical patterns, such as noun phrases, separately from the rest of human behavior. This chapter looks at a collection of cases in Swedish, English, and Estonian, where the speaker initiates a noun phrase but completes it with an embodied demonstration. Other participants treat this multimodal structure as complete and comprehensible. Building on earlier research on syntactic-bodily units (Keevallik 2013, 2017) this study calls into question the analytic boundary between language and the body and argues that grammatical projection cross-cuts modalities even within the assumedly robust “noun phrase”.

¹ This research was carried out within the project “Vocal Coordination of Human Action”, funded by the Swedish Research Council 2016-00827.

Keywords: syntactic-bodily units, multimodal “noun phrase”, Swedish, English, Estonian, dance instruction, interactional linguistics, determiner, preposition, embodiment

Introduction

In contrast to non-temporal and hierarchical models of grammar, a growing body of linguistic research acknowledges the fact that grammatical structure emerges in real-time interaction and may even be considered an epiphenomenon of what people do when engaging with each other (Hopper 1998). Linguistic units are accordingly being re-conceptualized as essentially temporal. Two crucial features of a real-time linguistic unit are its projectability and its completion, as recognizable for the participants. Projection reflects the temporal nature of spoken interaction and means that an individual action, or part of it, foreshadows another one in the near future (Auer 2005, Streeck & Jordan 2009), such as an article that projects a noun and a story preface that projects a story. Projectability seems to be an inherent characteristic of the linguistic-vocal mode of human action, while the realization of the projection can be embodied, resulting in *embodied completion* (Olsher 2004, Mori & Hayashi 2006). The occurrence of so-called *syntactic-bodily units* (Keevallik 2013, 2017) forces us to rethink the

nature of syntax in terms of its cross-modal capacity of projection. Units in interaction are not always accomplished exclusively within the vocal channel.

Human beings constitute meaning and action in a way that integrates speech and the body (Goodwin 2000, 2002, Jones & LeBaron 2002, Schegloff 2005, Stivers & Sidnell 2005, Enfield 2009, Mondada 2014). Meanings can, for example, be achieved by combining the verbal mode of communication with various kinds of gesture, and the gestures may contain absolutely crucial information for the participants. When it comes to gesture, Slama-Cazacu (1976) talked about *mixed syntax*, Enfield (2009) called the resulting units *composite utterances*, and Goodwin's (2006) term is *multimodal utterance*. However, since uttering involves speech, the term *utterance* may not be the best one to characterize instances when speech is not clearly profiled. It may convey the assumption that speech is somehow still most crucial for communication, even though in several human activities this is clearly not true. Reference to *syntax* involves the same limitations. Slama-Cazacu (1976: 216-222) talks about kinesic elements "replacing" syntactic elements, thereby treating language as primary in relation to other semiotic devices. The primacy of language, however, has increasingly been questioned in recent years (Mondada 2014). A neutral term in this sense would be *hybrid construction*, coined by Mori and

Hayashi (2006) for characterizing interactional moves involving both talk and gesture in conversation between L1 and L2 speakers, but *construction* is also a variously defined term in contemporary linguistics. Instead, the current study uses the term *syntactic-bodily unit*, building on the conversation analytic tradition (especially the programmatic article by Sacks et al. 1974), and focusing on projection and completion in particular. In contrast with the above studies it looks at whole-body demonstrations along with gestures and discusses their temporal matchings with speech in the activity of dance instruction. It targets verbally initiated noun phrases that are completed in the embodied mode, achieving a meaningful complete action in an inherently multimodal manner. Furthermore, as is common in real-time interaction (Auer 2009), these syntactic structures can then be built on further, adding lexical or clausal material. Accounting for embodied demonstrations as sense-making elements will among other things have consequences for how we analyze central conversational mechanisms, such as repair and reformulation.

Noun phrase in this study is understood in a traditional manner as consisting of a noun as its head and various non-head elements, including articles, demonstrative determiners, and adjectives (Payne 1992: 33-34, 86; Genetti 2015: 121-124). The study is not a contrastive one, but does build on materials in Swedish and English, which make use of articles and

prepositions, as well as Estonian, which does not and where the options of projection within a noun phrase are therefore fewer. The aim is to put forward a thoroughly multimodal and temporally evolving account of noun phrases, showing that they are a result of interactional negotiation, perhaps even an epiphenomenon. Similar arguments have been made for spoken noun phrases in Finnish interaction by Helasvuo (2001). It is remarkable, though, that in earlier studies on embodied completion (Weeks 1996, Olsner 2004, Mori & Hayashi 2006, Haviland 2007), there are no examples of verbally initiated noun phrases. This may of course be an incidental result of the fact that these authors had no particular focus on the syntactic details. It is also possible, though, that multimodal noun phrases are more likely to appear in contexts where the bodies are being instructed, such as in dance classes which will be explored in the present paper.

The data and the collection

The main body of the data comes from 29 hours of video recorded group dance classes teaching lindy hop and balboa in Sweden. The groups range from 6 to 60 students and there are altogether 14 teachers in the recordings. Six teachers speak Swedish (13h), and ten English (15h). One teacher couple gives classes in both Swedish and English, and there is only one

native speaker in the English data. Swing dances are often taught by a teacher couple, in which the complementary dancer roles are called *lead* and *follow*, and I have chosen to use this endemic referring practice in this paper. All the teachers have signed a formal consent to use the recordings for research and all the students have agreed to it. The complex actions in focus in the current study are mostly accomplished during extended teacher turns, the so-called instructive segments of the dance classes (Broth & Keevallik 2014). These are segments when practice has been stopped, the students are standing and watching the teacher(s). During these turns the teachers may deal with a number of pedagogical issues and the emerging syntactic-bodily structures may be either corrective or presenting new tasks for practice.

In addition, 9 hours of Estonian dance classes with 3 different teachers were recorded in Estonia, but merely one noun phrase completed by embodiment was identified in these data. There may be two reasons for that, a structural and a cultural one. Differently from Germanic languages, Estonian does not have grammatical articles. At the same time, words such as *üks* ‘one’ and *see* ‘this’ have been shown to mark indefiniteness and recognizability, respectively (Pajusalu 1997, 2005). These are used for projection in conversation (Keevallik 2010, 2011), and could in principle show up before embodied demonstrations in the data. However, as can be judged from the

recordings, there are a range of differences between the pedagogical cultures in Estonia and Sweden. In the Swedish settings the teachers tend to talk and demonstrate simultaneously during instructive segments of the class. In contrast, in the Estonian classes the teacher mostly stands on the side and shouts out her instructions and comments to the students. As the practice described here involves the simultaneous deployment of language and the body, the Estonian teachers in the recordings do not use it as frequently.

The collection includes 28 cases where an initiation of a noun phrase (NP) is followed by an embodied demonstration, and then optionally by another verbal element in the NP. There are 12 instances in English, 15 in Swedish, and one in Estonian.² For the sake of a clear delimitation of the collection, the cases where the embodied demonstration is accompanied by step names or beat numbers were excluded, as they would verbally complete the noun phrase. Their temporal organization may be identical to the focus cases, where the demonstration is either accomplished in silence or accompanied by non-lexical vocalizations. About a quarter of the instances (6) do not involve vocalizations produced in synchrony with the demonstrations, and these short periods of silence are spread across all the different NP patterns. Importantly, no cases were included where the descriptive noun was uttered

² The dearth of "multimodal" NPs in Estonian may only partly be explained with fewer projective resources (lack of articles), as the teachers were generally less active in demonstrating the dance moves.

after the embodied demonstration, as that would have just resulted in a regular noun phrase, at least retrospectively. An example would be *what we usually use in balboa is a lot of step hold*, where the “step hold” is demonstrated simultaneously. This type of cases were excluded. The focus here is on the fact that there does not need to be a noun produced as a kind of a “replacement” for the bodily demonstration at all.

The analysis will start from instances where there is just one item produced in a noun phrase before the embodied demonstration, moving on to instances with several items, showing that these multimodal units are not limited to particular grammatical structures. As a note of caution, this explorative study cannot aim at an exclusive collection of all possible patterns in the three languages, nor does it systematically pursue the contrastive affordances among the languages. Rather, it will present a first overview of the variety of syntactic-bodily structures within an NP in a dance teaching context.

Simple NP-initiations: an article, a demonstrative, an adjective

Article

To start with, there are five instances in the data where the dance teacher produces an article within a clause in progress but then goes on to complete the unit with an embodied demonstration. In four cases in English, the definite article *the* occurs before the demonstration. In example 1, the teacher of the lead role in the couple (TeaL, the male, with his back to the camera) is explaining the correct relaxed feeling when doing a step called *swingout*. He initiates a clause *you feel like there is the* but does not utter the projected noun. Instead, he immediately launches a demonstration of the “laidback attitude” with his body, accompanied by and accentuated with vocalizations. The projecting syntactic structures are marked in bold in the examples, the preparatory phase necessary to begin dancing (e.g., bending the knees, lifting partner’s hand) is marked with a ^ and the embodied demonstration with an asterisk.

(1) Definite article, English (file: there is the)

```

1 TeaL:      so just to- for you to have that (.) little LAIDback,
2 teal&f      (0.5) attitude. when you swingout. that you ^feel
               teal&f                                     ^prep
3 teal&f      #like there is the *DA:# dah (0.5) DA: da dah dah (.)
fig           #fig.1                                #fig.2
4 teal&f      bahoup and (0.3) DA: dah (0.4) d-a: da
               -->

```

1. "like" - preparation

2. "DA:" - first dance step

Importantly for the timing, the lead teacher takes hold of his partner, the follow teacher, long before the actual demonstration, on the word *that* in line 1 of the transcript. At that moment he already projects an imminent couple dance demonstration. On the word *feel* in line 2 he begins the preparation for the dance (slightly lifting his left hand and bending his knees, figure 1), which he launches immediately on the beat after the article *the*. The embodied demonstration is thus perfectly well-timed for what the article projects. Furthermore, none of the students display any signs of understanding after the verbal noun phrase is discontinued. They do not treat the contribution as complete yet.

Likewise, the combination of a preposition and article may be followed by an embodied completion. A case is shown in example 2, where a teacher

suggests to her partner (as well as the students) a specific place in the choreography to start the practice. She projects her embodied demonstration with the NP initiation *from the*. On *the* she bends her knees in preparation for the first dance step on *zak*.

(2) Preposition + definite article, English (file: From the)

```
1 TeaF:      Fro:m ^the:: *zak (0.4) uhm (0.4) pa:: pu pah (0.3)
   teaf              ^prep *demo starts -->

2           zhidegi pum zet tet tet (0.9)
3 TeaL:      yeah. (.) * sure.
   teaf              demo ends*
```

The follow teacher's presentation is simultaneously an inquiry to her co-teacher TeaL (she keeps her gaze toward him throughout her dance) and an instruction for the students as to what exactly they should start practicing, once the music starts. Importantly, her demonstration continues until the co-teacher verbally acknowledges that he has identified the suggested moment in the choreography and agrees to this being the moment to start the practice with *yeah* in line 3. To underline, nobody treats TeaF's contribution as complete after the verbal production is discontinued after *from the*. The first reaction is by TeaL at the very moment when the dance demonstration has arrived at the point when it is recognizable as one specific pattern in the choreography. The grammatical projection is only treated as constituting a comprehensible whole together with the embodied performance that ensues.

In the first two examples we have thus seen a similar temporal pattern of a transition from the verbal to the embodied, while the embodied segment is accompanied by vocalizations. At the same time, there is a difference in the character of the demonstration between the two cases. While example (1) illustrated the style of dancing, example (2) aimed at referring to a particular dance pattern, which could only be identified after a number of beats into the dance. Both demonstrations started from beat “one”, but in example (1) the relevant information was provided immediately, while in example (2) it took a while before it could be performed (approximately at *tet tet* in line 2). It is as if the temporal dance pattern has to be scrolled from its conventional beginning to the correct place in front of the participants’ eyes. These two different options recur in the current collection.

In verbal-only activity contexts the projection of a definite or indefinite article would be fulfilled with a noun in the language. However, in the dance class context, or perhaps any context where the body is being instructed, there is a constant need to refer to the body and its movements of various kinds. This context thus reveals the capacity of grammar to also mobilize the body to create meaning in interaction. Crucially for the focus of this volume, the body and its sounds may be timed to complete a traditional “noun phrase”. The temporal fitting of the bodily demonstration is similar to the production of a noun, which is projected to follow

immediately after the article. Indeed, this also happens in the dance class data: an article may be followed by a noun. In addition, a noun that refers to a move or a step may be simultaneously accompanied by the bodily move, such as *from here we are gonna do the back step*, where the back step is actually performed during the last two words. This, among other things, illustrates the projective power of articles: we interpret what comes after them as a noun, if we can, even though it may as well be an accompaniment to the performing body. The above multimodal units are thus closely related to many other ways of coordinating the language and the body in dance instruction.

Demonstrative

In eleven cases the embodied demonstration was initiated by a demonstrative determiner and here we have examples in all three languages. Apart from when the demonstrative has a special determiner format, such as the Swedish *denhåra* and *detdåra* (vs. *den* and *det*), it is not necessarily easy to distinguish the determiner vs. pronominal usage of a demonstrative. Relying on prosody (continuing vs. terminal) and the subsequent grammatical development of talk, only instances as clear as possible were included in the collection.

Here is an example (3) in Swedish, where the teacher dances and simultaneously comments on the problematic parts of student performance (lines 1-3). In line 4 she has stopped dancing and is standing still, while explaining further. In line 5, similar to the above examples, the teacher utters a demonstrative determiner *dehåra*, approx. ‘this’, that projects a noun, but instead of a conventional word she engages her hand at this very moment to show what she means. In addition, she reuses the vocalization *pum* from her dance demonstration (lines 1,3), where she uttered it to mark the correct bouncy steps. Overall, this is an instruction to keep a distinct rhythm throughout the step pattern. The teacher asks the students to keep bouncing and throughout lines 5-6 she also displays the bounce with her own body. The initiated NP in line 5, however, is completed by iconic hand gestures that are produced in synchrony with the vocalizations *pum pum* and represent the bounces.

(3) Demonstrative, Swedish (file: pum yeah)

```

1 Teacher:  PUM. hh yeah:. (.)
2           å här blir det slapt, (0.2)
           'And here it becomes sloppy'
3           PU:M. hh yeah,
4           å- å det tycker jag är synd, att
           'And I think this is unfortunate that/so'
5           behåll dehåra *PU:m (.) PU:m*
           keep:IMP this
           'Keep this'
           tea * beat gestures*

6           det ska bara *bumma dunså PÅ* på nåt sätt.
           'It should just keep bumming bouncing somehow.'
           tea *beat gestures *
```

As can also be seen in the current example, these multimodal “noun phrases” tend to emerge after longer explanations, when the context of the demonstrated move has been set, which supports the sense-making of them in real time. In the current instance the teacher furthermore continues with a reformulation of the same idea immediately after the embodied demonstration is complete. The utterance *behåll dehära PUm PUm* ‘keep this PUm PUm’ together with the gesture (line 5) constitutes a comprehensible and coherent whole in itself, which allows the speaker to launch a next one after it without a hitch. She chooses to reformulate the same idea in another complete utterance *det ska bara bumma dunså PÅ på nåt sätt* ‘it should just keep bumming bouncing somehow’, this time with canonical Swedish syntax (interestingly inventing a verb *bumma*, approx. ‘to bum’ on the spot, line 6).

The only Estonian case in the collection involves the demonstrative *see* ‘this’ that can occasionally display features of a definite article (Pajusalu 1997). It occurs soon after the students have completed a dance routine and the teacher has asked whether they needed help remembering some of the steps. A student takes the chance to ask a question (example 4). He produces an initiation of a clause and, arguably, an initiation of a noun phrase with the demonstrative *see*. He stands arms akimbo during the talk in line 1 (as shown in figure 3) and demonstrates the step precisely during the

pause and the vocalization *km* (as shown in figure 4). During line 2 he is slightly swaying and gesturing, but not really showing the step any more.

(4) Demonstrative, Estonian (file: mis on see)

1 Student: SEE üks asi mis on see::# *(0.7) km# q*=
 'this one thing that is this'
 stu * demo *
 fig #fig.3 #fig.4

2 =ja siis oli vahepeal seal mingi jalgade vahetus.
 'and then in the middle there was a change of feet.'

3. Akimbo position

4. "km" - the dance step

See 'this' with level intonation projects a continuation (Keevallik 2010) and potentially an embodied demonstration. The incomplete intonation suggests that in this particular case it is used as a determiner. Furthermore, the student continues unproblematically (no repair or restart) to build on the multimodal unit with another coordinated clause in line 2. The embodied demonstration of the dance step is perfectly timed with the pause and the

brief vocalization, so that the projective determiner together with the demonstration emerge as a coherent multimodal structure. The body movements are here entirely subordinated to talk and its pace; that is, the speaker moves his body only at the very moment when a reference to a step is due and stands still the rest of the time before and after the embodied demonstration. The determining demonstrative *see* initiates an NP that ends up being completed by the body instead of a lexical noun. The speaker then smoothly continues with another monomodal utterance (line 2) that is tied to the previous multimodal one with the conjunction *ja* ‘and’.

As was the case with the articles above, a preposition plus demonstrative may be used for projecting an embodied completion. This is shown in example 5 where the teacher couple continues dancing while the lead teacher talks and vocalizes. Thus, in contrast to example 4, he is subordinating his talk to the dance tempo (see Keevallik 2015 for a discussion of these different options). The demonstrative *that* used here is a determiner that projects a reference to a moment where the dancers come out of a certain move in order to make a transition to the next one. The segment that the teacher is referring to with *out of that* occurs during line 2, and the word *ONE* already accompanies the ensuing step pattern (line 3) that illustrates how to come out of the move referred to with *that*.

(5) Preposition + demonstrative, English (file: out of that)

1 Teacher: so we come out of that (.)
 2 *uu puu puh (.) chuck (0.3)*
 tea * demo of referred steps *
 3 **ONE two three (.) five six seven eight**
 tea dance continues --->

In fact, it is even problematic to mark a specific point in the transcript where the reference would occur since the idea is to talk about the transition to a next dance step and the referred moment could as well be the very end of the step sequence, rather than all of it. This kind of indeterminacy in the exact location or length of the embodied referent is characteristic of the bodily mode of communication, as opposed to, e.g., words in a language. It takes expert knowledge by the participants to figure out the meaning insofar as they need it to improve their performance of that particular move.

Adjective

There are six examples of an adjective launching an NP in the collection, with or without a preposition, where the projection is fulfilled by the body. In example 6 the teacher, who is currently standing still, explains the right amount of pressure between the partners: it has to be short and rhythmical. In line 2 she initiates a clause and a noun phrase with the adjective *långt* ‘long’ but instead of a noun she brings her two arms together into a demonstration of excessive pressure, accompanied by an extended strain grunt.

(6) Adjective, Swedish (file: SUBS finns inget långt)

```
1 Teacher:  för det finns inget - (.)  
            'Because there is no -'  
2          det finns inget ^långt .h *kghkghgee*  
            'there is no long'  
            tea                ^prep      *gesture  *  
3          utan det finns ba *MM tjaka BUM tjaka BUM  
            'but there is only'  
            tea                *gesturing -->
```

The multimodal “noun phrase” in line 2 is followed by a contrastive statement about what the contact should be like instead, similarly accompanied by a gestural demonstration – but this time no noun phrase is initiated. This example, yet again, illustrates an instance where the embodied demonstration is subordinate to the temporality and logic of talk: the teacher starts talking and only at the moment when it becomes relevant in talk does she deploy her body for a demonstration. The completeness of the multimodal unit in line 2 is underlined by the fact that the speaker unproblematically continues to build on it with the grammatically contrastive *utan* ‘but’-structure in line 3. It would be impossible to provide an adequate (participant) analysis of the syntactic structure here without fully recognizing the legitimacy of embodied demonstrations in providing central content (see Keevallik 2017 on multimodal contrastive structures in dance classes, including the analysis of the Swedish *utan* ‘but’).

An adjective may also follow a preposition. This happens in the ensuing excerpt (example 7) where the Swedish adjective *olika* ‘various’ is uttered

after the preposition *i* ‘in’, before the embodied demonstration is launched. The instruction is here about variations of the upper body during a particular step called *Suzie-Q*. Accordingly, the teacher is continuously doing the same step with his legs, while changing the position of his torso and hands. He starts performing Suzie-Q together with his partner on the word *hamnar* ‘ends up’ in line 1, but it is only after the talk is abandoned that he raises his shoulders and curves his back into a demonstration of a specific styling. The stepping also simultaneously changes its character from neutral to tight and sharp, resulting in a distinct demonstration of a type of style. The follow teacher moves accordingly from neutral to sharp, mirroring the lead teacher, who is talking.

(7) Preposition + adjective, Swedish (file: hamnar man i olika)

```
1 Teal:    när man: nu dansar här, så (.) ^hamnar man i olika
           teal                                ^steps start
           'When you dance here you ends up in various'

2          * (3.0)
           teal&f    *demo -->

3          det här är också Suzie-Q.*
           teal&f    'This is also a Suzie-Q'
           demo continues --> *
```

Perhaps interestingly, this is the only instance in the data when a plural noun is projected by the initiation of a noun phrase. There is thus no restriction with regard to whether the projection is in singular or plural; the body can take care of both alternatives and complete the unit. The teacher actually presents a single movement option during the pause and continues displaying that during the comment in line 3. He formulates what he is

currently doing with his body, namely a stylized version of the step. The comment treats his own body as a relevant source of information in the ongoing activity and at the same time marks the completion of the structure launched by the NP-initiation. The temporality of the body is quite complex, as the teacher couple is actually dancing all the time, changing their positions along the way. Merely one clear switch in the dance style is temporally fitted after the initiation of an NP.

After this overview of three linguistic categories in an NP that can immediately precede an embodied demonstration, article, demonstrative and adjective, let us look at some combinations of these.

Complex NP-initiations

As it happens, a majority of the NP-initiations in the data were simple, but there were also two possible combinations of items, both including an adjective. In three instances the initiation consisted of an article + adjective and six instances contained a demonstrative + adjective. Characteristically, all of the adjectives were diminutives *little*, *a bit of a*, and *lilla*, *liten* 'Swe. small', so we can actually see this as a set of relatively fixed expressions

where the diminutive adjectives project either a verbal noun or an embodied demonstration. In the following a selection of examples will be presented.

Article + adjective(s)

A simple case with an iconic gesture can illustrate a multimodal pattern where an article is followed by an adjective, resulting in a projection which is fulfilled by the body. In example 8 the teachers have been demonstrating a choreographed step sequence and now one of them is summarizing it with a multimodal presentation. He initiates a copula clause and an NP with an indefinite article *a* and two consecutive adjectives *nice* and *little* (line 1). The gesture that depicts the movement of the bodies through the space during the step pattern is launched already during the verbal initiation of the clause and then accompanied by repetitive vocalizations throughout. In the transcript the gesture is marked from the moment when the hands are in position for the gesture until they are lowered all the way down, back to a possible home position. Figures 5-8 show different stages of the gesture in relation to vocal production, they show first a lateral movement in space (figure 6) and then a back-and-forth movement (which becomes up-and-down in the gesture, figure 7).

(8) Article + adjective, English (file: nice little X)

```
1 Teal:      so yeah, it*'s #just a ni#ce little
   teal          *gesture -->
```

```

fig                                #fig.5    #fig.6
2      ji#ggiiggijiggi.* #
teal    -->*
fig      #fig.7          #fig.8
3 TeaF:  jiggijiggijiggi.

```

5. Hands prepared

6. "nice" - hands open and back

7. "ji-" - hands up and down

8. After completion

Importantly, the acknowledging repetition of the vocalization *jiggijiggijiggi* by the follow teacher in line 3 treats the multimodal structure in line 1 as complete not after the verbal part is discontinued (i.e. after *little*) but when

the embodied demonstration is terminated. Her turn follows precisely after TeaL has uttered *jiggijiggijiggi*, lowered his voice into a prosodic completion and moved his hands back to home position (end of line 2, shown in figure 8). We thus have emic evidence of these multimodal structures being treated as real and coherent in this activity setting.

A similar case preceded by a preposition is shown in example 9. The teacher is explaining how to finish off a step sequence with an eye-catching improvisation. He projects this flashy demonstration with a combination of preposition, article and adjective, *with a litte* (line 1). He is, however, already dancing when he initiates the clause, so the beginning of the actual demonstration is somewhat difficult to distinguish. Clearly, though, he moves from neutral and rhythmic dancing to (mock) flashy and non-rhythmic dancing right after the talk is abandoned, which is marked as the “demo” in the transcript. Similar to example 7 we here see a comment on the performed dance move after the demonstration is complete (*that’s cool* in line 2), treating the embodied part as meaningful in the context.

(9) Preposition + article + adjective, English (file: with a little)

1 Teacher: and you finish it off with a little *(2.0)*
tea *demo *
2 tha*t's cool, isn't it.

The demonstration is again temporally fitted to the verbal initiation of a clause, appearing at the precise moment when a noun could have been

uttered. After the completion of the dance step the teacher goes on to comment on his own performance with another clause. At that moment he also stops dancing and turns to the students for an appreciation of his pedagogical point which could only be inferred from the combination of verbal and embodied modalities of his contribution.

Demonstrative + adjective

The most frequent combination in the collection is a demonstrative together with an adjective, and that can occur as demonstrative first, as in *this little*, or adjective first, as in *liten sänhåra* ‘Swe. little suchlike’. A case is shown in example 10 where the teacher has made clear that the students are performing the steps too heavy: they are “working” too hard. In the excerpt the teacher, who is standing still and talking, guides the students to make the steps lighter. Line 3 is a reformulation of line 2, and at the moment when she has uttered the attributive *den här* ‘this’ and the adjective *lilla* ‘little’, projecting a noun, she launches an embodied demonstration.

(10) Demonstrative + adjective, Swedish (file: SUBS det här lilla 2)

- 1 Teacher: utan försÖK nu:, en gång, (0.3)
'but TRY now, once,'
- 2 att TÄNka att det inte ÄR så mycket arbete,
'to THINK that it IS not so much work,'
- 3 att det ÄR bara den här ^lilla
'that it IS only this little'
tea ^prep

4 *du:bidu pu pu tʃagidu PA,
tea *demo starts -->
 ((demo goes on for three more "eights")

In all three of the languages of this study, the determiner and pronominal uses of demonstratives may be difficult to disentangle, but in combination with an adjective, as above, they can more clearly be analyzed as projecting a noun. The vocalizations that follow, however, can in no way be categorized as words in a language, as they are *ad hoc* variable sounds that belong closely together with the dance steps performed (Keevallik 2015: 317-320). In this episode the teacher prepares the dance steps slightly before the verbal structure is abandoned so that the beginning of the actual dance step can again be exactly fitted to the end of the last word. As in several of the examples above, beat “one” in the dance is synchronous with the first vocalization syllable *du*. Since the argument she is making concerns the overall style of dancing, the relevant information is presented from the very beginning of her embodied demonstration, and she could basically stop her demonstration after a few beats. Nevertheless, she continues to perform the dance for quite a while, speeding up the moves and illustrating how “small” they can be performed. This again shows how the body is organized differently from talk, enabling temporally extended demonstrations after the initiation of a mere noun phrase. In this example, it is as if the logic of the dance starts to predominate after getting into the demonstration. The dance lasts for three conventional patterns, counted in “eights” in lindy hop, which constitutes a contribution that is easily understood by the participants in this

activity setting. Examples like this with extended performances make it especially clear that embodied demonstrations cannot be just squeezed into the grammatical category of nouns but constitute a qualitatively different mode of communication, albeit still projected by, and timed with, canonical initiations of noun phrases.

To summarize, in some activity contexts we can see that the body is timed to produce essential content right at the moment when an NP has been initiated. As an alternative to silence or vocalization, the bodily demonstration may be accompanied by a term, a step name or a count number, i.e. a noun. In that case the projection would be completed with linguistic devices. In contrast to a noun phrase *par excellence*, however, there is no clear limit to when a multimodal unit is complete. A dance demonstration can go on for a considerable time, and its length may rather be defined by the structure of the dance than the logic of talk. The multimodal practices are furthermore in the service of the pedagogical aims of the teacher. Especially for these reasons it would be very difficult to maintain that there is something specifically noun-like happening in the demonstrations. Among other things, nouns have been claimed to typically express time-stable entities and here we have an evolving movement (see, however, about the relativity of this criterion in Payne 1997: 33). Crucially, the same kinds of demonstrations are performed after a variety of

grammatically projective structures, such as initiations of verb phrases, prepositional phrases and contrastive conjunctions (Keevallik 2013, 2015). In short, the embodied demonstrations are a different kind of entity, as compared to words in a language, but they can nevertheless fulfill a grammatical projection and constitute a meaningful unit together with it.

Participant orientation to a multimodal “noun phrase”: Collaborative completion

In order to ultimately demonstrate that a linguistic-bodily structure is real for the interactants one could document whether and how the participants orient to it, such as starting to talk at a particular moment in time. An orientation to incompleteness of a unit may, for example, imply that other participants refrain from talk until it is terminated. This is what we have seen across the examples presented so far. Another way to display understanding of incompleteness in talk is to provide the completion for someone else's initiation (Lerner 1995, Hayashi 2005). In that way two participants co-construct a unit and simultaneously display a mutual understanding of what this unit should look like. There is one instance in the current collection where a teacher utters the verbal initiation of the noun

phrase and the other teacher completes it with a movement and a vocalization.

In example 11 the teacher couple is jointly involved in a presentation of a dance move. The lead teacher (TeaL) is talking and preparing his partner for a spin. The spin is timed precisely with the end of his projective segment of talk, consisting of a combination of a Swedish preposition *i* ‘in’, an indefinite article *en* ‘a’, an adjective *liten* ‘little’, and a colloquial determiner *sånhåra* ‘suchlike’. This initiation projects a continuation partly through the incomplete grammar, partly through the specific determiner form of the demonstrative *sånhåra*, and partly through the completely level prosody on *sånhåra*, which suggests no completion. In terms of content the projected unit needs to refer to the move of the follow, because the lead does not “end up” anywhere, he guides the follow, and it is the follow who ends up spinning. During the spin the teacher performing the role of a follow in the dance (TeaF) produces a vocalization *wii*. Together with the movement that fulfills the grammatical projection launched by the lead teacher. The multimodal “noun phrase” thus emerges across two different speakers and entails a verbal projection followed by an embodied completion accompanied by a vocalization.

(11) Co-production, Swedish (file: Ljud17wii1)

1 TeaL: så att man hamnar i (.) en (0.3) liten sånhåra
 ‘So that one ends up in a little suchlike’

```

2          *(1.0)
  teaf&l    *demo -->

3 TeaF:    wi#i:.*
           -->*
  fig      #fig.9

4 Students: khh heh hehe hahah

```

9. Mid-spin on "wi-"

Differently from all other vocalizations in the current collection, *wii* is half-conventionalized as a syllable used to accompany exciting movement, such as when swinging a child. Likewise, it is here used to underline the pleasant feeling of a spinning body for the benefit of a pedagogical presentation of a dance move. It is produced at the moment when the lead has just set the follow's body in rotation and lets go of it, in synchrony with the follow's

own airy feeling of being in spinning motion following through on the impetus. The *wii* ends when the rotation ends and the partners face each other. Crucially, the students immediately react to the termination of the spin and the *wii* by collective laughter, which shows their orientation towards the unit as complete. One could, on the other hand, ask whether the *wii* might be a response rather than a continuation of the structure launched by TeaL. Clearly, it could not be responsive to what he said, because line 11 is a description of the couple's current moves directed to the students, not to his partner. It could only be responsive to what he does with his body but, more precisely, it is an externalization of what TeaF's own body is currently experiencing; an intense rotation. The timing of *wii* reflects her own movement, not his. Furthermore, just like TeaL's talk, her vocalization can be argued to be done for the onlooking audience, and thereby emerges as building one whole with the TeaL's projection. Last, TeaL does not make any attempt to complete or re-complete what he left unfinished, treating the spin + *wii* as a sufficient completion. The two teachers behave as one party during this couple dance class, co-producing pedagogical points and the syntactic-bodily unit *man hamnar I en liten sânhâra wii* 'one ends up in a little suchlike wii' is accordingly achieved across two different participants.

Verbal completion of a multimodal "noun phrase"

After a demonstration, the speaker-performer can continue to build the grammatical structure they have initiated. There are two instances in the current collection where the embodied element emerges as if in the middle of a “noun phrase”, as a grammatical modifier. A case in point is shown in example 12 where the teacher accomplishes a brief demonstration of a lower body rotation during the vocalization *quh* (figure 11) but then goes on to complete the unit with the word *feeling*, which means that the movement + *quh* emerges as a kind of an adjective in this structure, between the demonstrative and the noun. The body is in neutral position before and after the short demonstration (figures 10 and 12).

(12) Embodiment as modifier, English (file: ATTR1feel)

```

1 Teacher:  uh, HERE, when we end up here, we are more
2           parallel to each other. ^so# it's not this (0.2)
   tea                                     ^prep
   fig                                     #fig.10

3           *↑QUH#*  fee#ling, .h^it's more (.) *side to (.) side.
   tea      *demo *                                     ^prep      *demo -->
   fig      #fig.11 #fig.12

```

10. preparation

11. "quh" - demo

12. Back to neutral

The embodied demonstration shows what kind of a feeling the teacher wants to address as problematic in the class. The vocalization *quh* that is precision-timed with the move expresses the bodily feeling through its shortness, sharpness and character (the higher pitch and the high vowel suggest a coquette move, as shown in figure 11). Accordingly, the demonstration provides crucial information for the students. Pragmatically the language used is in the supportive role, containing generic words such as *feeling*, while the demonstration and the vocalization *quh* (with a glottal stop in the syllable onset) provide a more fully specified information of

what the dance movement should look and feel like. It may be the case that if the NP is to be continued the demonstration has to be accompanied by a vocalization. More generally, the existence of these additional patterns, such as a verbal completion after a multimodal structure, further illustrates the way language and the body are jointly involved in the process of constituting meaningful “NPs” in interaction without any hitches in production or displayed problems with understanding. Examples like this witness of the totally smooth incorporation of multimodal units into the syntagmatic structure of talk.

Another way to display that the multimodal “NP” is non-problematically part of the sensemaking processes in a dance class is to continue with a next clause. This is the most common option in the collection: either the speaker herself or someone else, e.g., the co-teacher, continues with a new clause, without repair or any other hitch in production. As we have seen in many of the above examples, the teacher can, for example, comment on what her/his body was doing (examples 3,4,7,9), or the other teacher might respond (examples 2,8). In example 6, we also examined the option to continue with further clauses by grammatically tying them to the prior one that included a multimodal “NP”: We saw an instance of the incremental building of a complex clause with the *utan* ‘but/instead’- construction with *det finns inget*

långt .h kghkghgee utan det finns ba mm tjaka bum tjaka bum ‘There is no long .h kghkghgee but there is merely mm tjaka bum tjaka bum’.

Repair or re-completion?

The above arguments about phrases potentially being multimodal have consequences for conversational mechanisms such as repair. When we account for the body as a meaning-making device, we may have to reconsider structural analyses of situations where the relevant information is not only provided in the stream of speech. In example 13 it can be argued that the teacher first completes the projection with an embodied demonstration but then also utters a noun. The structure in line 4 that is initiated by *so we need to get this* could have been complete after the (0.7) pause during which the teacher keeps bouncing; we could perceive this as an entirely comprehensible and completed contribution in this activity setting, asking the students to perform pulse while they are turning. The teacher then, however, repeats a part of the already produced verbal structure, the demonstrative determiner *this*, and now completes the NP verbally with *pulse* (line 4). The verbal gloss *this pulse* thus emerges as an alternative version of what the body has already produced, and not a repair, as it appears if we only consider the verbal stream *this (0.7) this pulse*.

(13) Glossing the body, English (file: this x through y)

```
1 Teacher: (X) that you're turning, (0.4) that would (0.3)
2          also be:: (.) super hard if you keep (0.6)
3          your foot flat on the ground the whole time.
4          *so we need to get this# (0.7) this pulse going,*.
```

* demo of bouncing *

As it happens, in this case the embodied demonstration has already emerged from the very beginning of the clause and the teacher also continues to bounce all the way through the ensuing NP. We thus cannot see a temporally neat fitting of the grammatical projection with an embodied demonstration here, but, as in the above cases, there is a temporal slot of silence in line 4 during which the body continuously demonstrates what the students “need to get”. One could thus consider these kinds of instances re-completions, as glosses provided to what was already shown by the body. Here the body has already illustrated the bouncing before the verbal expression *this pulse going* (at the end of line 4) provides a verbal gloss for it.

To summarize, it is crucial not to ignore the embodied demonstrations that take place in a temporal slot in these complex turns. They provide essential content and crucially participate in achieving intersubjectivity between the participants in this activity.

Conclusion

This paper has illustrated what a study of real-life language use can reveal about one of its central categories, the NP. Among other things, the flow of speech may be intertwined with embodied demonstrations that are precision-timed with talk to build meaningful multimodal “noun phrases” in real time. It was shown that an embodied demonstration may be tightly fitted after an article, a demonstrative, an adjective, or a combination of those, and that the participants orient to completion not after the projecting grammar is discontinued, but after the embodied demonstrations make sense for the practical purposes at hand. After the embodied demonstration, the speaker-performer can non-problematically continue building the NP or a more complex clausal structure, or they can just produce a next clause. An analysis that incorporates the contribution of the body influences how we analyze conversational mechanisms, such as repair, as it may happen that the body has already provided essential information before it is uttered in a stream of speech.

The patterns in focus in our study are thus not entirely different from other emergent multimodal structures in a dance class. In many cases regular noun phrases are launched and completed with a noun, while in others the meaning is conveyed without the noun, using the body instead. We have

thus explored very specific versions of a “noun phrase” in embodied practices, showing that it is not always literally a noun that appears as the head in it. At the same time, the initiating segments involve elements that have been recognized in verbal noun phrases, suggesting a striking regularity across modalities where the core content can be conveyed either by a noun or by the body. They could be analogs of arguments as well as obliques, just like regular noun phrases. We might wonder, though, whether the resulting patterns still qualify as noun phrases, or whether we should establish another category for these multimodal structures. Characteristically, the verbal parts of such a structure would merely play a pragmatically supportive role in relation to the fully specified content displayed by the body.

The above analyses provides a richer account of what happens to “noun phrases” in interaction, transgressing modalities and activity types. While activities where bodies are being instructed may be an exceptionally prolific place for finding out how different communicative modalities are coordinated in time, it is by no means an exceptional way of using language (Mondada 2016, Keevallik 2018). For research on interaction, the contribution of the current study is a demonstration that units can consist of syntax and an embodiment in a manner that is very similar to that of verbally-built units (Li 2013). They display projection that is realized in a

completion but the two are done in different modalities. This paper thus makes yet another argument in support of the temporal nature of grammar and the need to study it as such, because this is how participants make sense of each other. “Noun phrases” are temporally evolving purposeful structures and grammatical projection works across modalities. It so happened that Estonian, a language without clear grammatical articles, did not display numerous occurrences of multimodal “NPs”. Further comparative work targeting the role of the body in building noun phrases in such languages, as well as in languages in which articles, adjectives, and demonstratives typically follow the noun, would be most fascinating.

Transcription conventions

The paper uses multimodal transcription conventions developed by Lorenza Mondada:

<https://www.lorenzamondada.net/multimodal-transcription>

References

- Auer, Peter. 2005. Projection in interaction and projection in grammar. *Text* 25(1): 7–36.
- Auer, Peter. 2009 On-line syntax: Thoughts on the temporality of spoken language. *Language Sciences* 31(1):1–13.
- Broth, Mathias & Leelo Keevallik. 2014. Getting ready to move as a couple: Accomplishing mobile formations in a dance class. *Space and Culture* 17(2): 107–121.
- Enfield, Nick. 2009. *The Anatomy of Meaning: Speech, Gesture, and Composite utterances*. Cambridge, UK: Cambridge University Press.
- Goodwin, Charles. 2000. Action and embodiment within situated human interaction. *Journal of Pragmatics* 32: 1489–1522.
[http://dx.doi.org/10.1016/S0378-2166\(99\)00096-X](http://dx.doi.org/10.1016/S0378-2166(99)00096-X)
- Goodwin, Charles. 2002. Time in action. *Current Anthropology* 43: 19–35.
- Goodwin, Charles. 2006. Human sociality as mutual orientation in a rich interactive environment: Multimodal utterances and pointing in aphasia. In *Roots of Human Sociality*, N. Enfield & S.C. Levinson (eds.), 96–125. London, England: Berg Press.

- Haviland, John B. 2007. Master speakers, master gesturers: a string quartet master class. In *Gesture and the Dynamic Dimension of Language: Essays in Honor of David McNeill*, S.D. Duncan, E.T. Levy & J. Cassell (eds.), 147–172. Amsterdam/Philadelphia: John Benjamins.
- Hayashi, Makoto. 2005. Joint turn construction through language and the body: Notes on embodiment in coordinated participation in situated activities. *Semiotica* 156: 21–53.
- Helasvuo, Marja-Liisa. 2001. Emerging syntax for interaction: Noun phrases and clauses as a syntactic resource for interaction. In *Studies in Interactional Linguistics*, M. Selting & E. Couper-Kuhlen (eds.), 25–50. Amsterdam/Philadelphia: John Benjamins.
- Hopper, Paul. 1998. Emergent Grammar. In *New Psychology of Language: Cognitive and Functional Approaches to Language Structure*, M. Tomasello (ed.), 155–175. Mahwah NJ: Lawrence Erlbaum.
- Jones, Stanley E. & Curtis D. LeBaron. 2002. Research on the relationship between verbal and nonverbal communication: Emerging integrations. *Journal of Communication* 52: 499–521.
- Jääskeläinen, Anni 2013. *Todisteena Äänen Kuva: Suomen Kielen Imitatiivikonstruktioit.* [Imitations of Sounds as Evidence: The constructions of imitatives in Finnish.] Helsingin yliopiston suomen kielen, suomalais-ugrilaisten ja pohjoismaisten kielten ja kirjallisuuksien laitos, Helsinki university.
- Genetti, Carol. 2015. *How Languages Work*. Cambridge: Cambridge University Press.
- Keevallik, Leelo. 2010. The interactional profile of a placeholder: The Estonian demonstrative *see*. In *Fillers, Pauses and Placeholders*, N. Amiridze, B. H. Davis & M. MacLagan (eds.), 139–172. Amsterdam/Philadelphia: John Benjamins.
- Keevallik, Leelo. 2011. Pro-forms as projective devices in interaction. *Discourse Processes* 48(6): 404–431.
- Keevallik, Leelo. 2013. The interdependence of bodily demonstrations and clausal syntax. *Research on Language and Social Interaction* 46(1): 1–21.
- Keevallik, Leelo. 2014. Having a ball: Immaterial objects in dance instruction. In *Interacting with Objects: Language, Materiality, and Social Activity*, M. Nevile, P. Haddington, T. Heinemann & M. Rauniomaa (eds.), 245–264. Amsterdam/Philadelphia: John Benjamins.
- Keevallik, Leelo. 2017. Linking performances: The temporality of contrastive grammar. In *Linking Clauses and Actions in Social Interaction*, R. Laury, M. Etelämäki & E. Couper-Kuhlen (eds.), 54–72. Helsinki: Finnish Literary Society.
- Lerner, Gene H. 1995. Turn design and the organization of participation in instructional activities. *Discourse Processes* 19: 111–131.
- Li, Xiaoting. 2013. Language and the body in the construction of units in Mandarin face-to-face interaction. In *Units of Talk -- Units of Action*, B. Szczepek Reed & G. Raymond (eds.), 343–375. Amsterdam/Philadelphia: John Benjamins.
- McNeill, David. 2005. *Gesture and Thought*. Chicago: Chicago University Press.

- Mondada, Lorenza. 2014. The local constitution of multimodal resources for social interaction. *Journal of Pragmatics* 65: 137–156.
<http://edoc.unibas.ch/dok/A6328898>
- Mondada, Lorenza. 2016. Challenges of multimodality: Language and the body in social interaction. *Journal of Sociolinguistics* 20(3): 336–366.
- Mori, Junko & Makoto Hayashi. 2006. The achievement of intersubjectivity through embodied completions: A study of interactions between first and second language speakers. *Applied Linguistics* 27: 195–219. <https://doi.org/10.1093/applin/aml014>
- Olsher, David. 2004. Talk and gesture: The embodied completion of sequential actions in spoken interaction. In *Second Language Conversations*, R. Gardner & J. Wagner (eds.), 221–245. London, England: Continuum.
- Pajusalu, Renate. 1997. Is there an article in (spoken) Estonian? In *Estonian Typological Studies II* [Tartu Ülikooli eesti keele õppetooli toimetised 8], M. Erelt (ed.), 146–177. Tartu: Tartu ülikooli kirjastus.
- Pajusalu, Renate. 2005. Anaphoric pronouns in spoken Estonian. Crossing the paradigms. In *Minimal Reference. The Use of Pronouns in Finnish and Estonian Discourse* [Studia Fennica Linguistica 12], R. Laury (ed.), 107–134. Helsinki: Finnish Literature Society.
- Payne, Thomas. 1997. *Describing Morphosyntax*. Cambridge: Cambridge University Press
- Sacks, Harvey, Emanuel A. Schegloff & Gail Jefferson. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50(4): 696–735.
- Slama-Cazacu, T. 1976. Nonverbal components in message sequence: Mixed syntax. In *Language and Man: Anthropological Issues*, W. McCormack & S. Wurm (eds.), 217–227. The Hague, The Netherlands: Mouton.
- Stivers, Tanya & Jack Sidnell. 2005. Multi-modal interaction. *Semiotica* 156: 1–20. [10.1515/semi.2005.2005.156.1](https://doi.org/10.1515/semi.2005.2005.156.1)
- Streeck, Jürgen & J. Scott Jordan (eds.). 2009. Projection and anticipation in embodied communication, Special issue. *Discourse Processes* 46.
- Teleman, Ulf, Staffan Hellberg & Erik Andersson (eds.). 1999. *Svenska Akademiens Grammatik* [Grammar of the Swedish Academy]. Nordstedt, Stockholm.
- Weeks, Peter. 1996. A rehearsal of a Beethoven passage: An analysis of correction talk. *Research on Language and Social Interaction* 29(3): 247–290.