Gestures, referents, and anaphoric linkage in learner varieties

Marianne Gullberg

1. Introduction

For discourse to be comprehensible, referents have to be uniquely identifiable. This paper is concerned with how gestures contribute to this process in interaction with speech in learner discourse. It is a common lay assumption that gestures are chiefly connected to exophoric, demonstrative expressions, such that gestures refer directly to and index entities in the real world. A typical example would be a speaker pointing to an object saying, "look at that". However, speakers' gestures also depict and index such abstract things as discourse referents and discourse. Gestures can be endophoric, and, more specifically, anaphoric. The systematic association between gestures, space, and speech allows cohesion to be established in all anaphoric domains. The purpose of this paper is to outline how such gestural cohesion can shed light on the specific characteristics of anaphoric linkage in learner varieties.

Target language-specific preferences for anaphora, cohesive patterns and information organisation in speech are acquired late, if at all, by adult language learners. Difficulties can be found in all anaphoric domains: space, time, and referential movement of first order entities. In the domain of referential movement, i.e. the movement of animate entities with a high degree of control over events and actions in discourse (Klein & Perdue 1997; von Stutterheim & Klein 1989) studies show that learner varieties at early non-finite stages of proficiency (the Basic Variety stage, Klein & Perdue 1997) typically rely heavily on implicit information established in the surrounding context. In contrast, at later stages learner varieties are *more* explicit than native varieties. In particular, the domain of maintained reference, in which agents typically move from topic to topic in successive utterances, is characterised by full NPs being favoured over pronouns and zero-anaphora even in contexts where referents are supposedly both accessible and recoverable (e.g. Ahrenholz 1998; Carroll et al. 2000; Extra, Strömqvist, & Broeder 1988; Givón 1984; Hendriks 2000; Muñoz 2000; Strömqvist & Day 1993).

(1)   *elle a eh + donné + pour eh la dame*
      'she has uh + given + to uh the lady
      et ehm + la dame allé pour une autre dame + eh +
      and uhm + the lady went to another lady'

In (1), a learner of French2 introduces the referent *la dame* in object position in the first utterance, and then uses the same NP in subject position in the immediately subsequent utterance instead of demoting the referent using a pronoun. While there may be pragmatic reasons in L1 for using a full NP to refer to an immediately preceding referent, learners seem to routinely use NPs in such contexts regardless of appropriateness.

The reasons for this over-extension of full nominal forms have been amply discussed. It has been suggested that, since lexical means are acquired before grammatical ones in all areas of acquisition, so lexical means of co-reference should be expected before more grammatical, pronominal ones. The pronominal forms of anaphora are also particularly complex as they encode multiple grammatical categories simultaneously, typically gender, number, and case. This complexity makes pronominal systems error-prone (Hendriks 2000). It has therefore been suggested that learners shun them in favour of full nominal expressions as a way of avoiding ambiguous reference created by erroneous pronominal forms (Williams 1988). Moreover, the choice of form depends on the intricate interplay between information organisation at a local level (e.g. given vs. old information) and at a global level beyond the single utterance (Carroll & Lambert, this volume; Carroll et al. 2000). This span over multiple levels clearly represents a challenge to learners whose processing capacities are already occupied by formal aspects at a local level, and may further motivate avoidance of pronouns. The influence of the first language (Jin 1994), of markedness factors (Muñoz 1995), and task (Muñoz 2000) have also been explored. In a more interactional perspective, it has been proposed that the use of full NPs is motivated by the social co-ordination of talk in progress. Full NPs can be used as a means to mark disagreement or manage viewpoint (Fox 1987; Pekarek 1999).

This paper will outline how gestural cohesion can shed light on the effects and causes of the specific characteristics of anaphoric linkage in spoken learner varieties. First, it will be shown that spoken learner varieties come with particular gestural profiles that are related to the characteristics of spoken varieties in non-trivial ways. Put differently, there is a learner variety of gestural anaphoric linkage which (a) differs systematically from the L1 variety in the same way the spoken learner variety differs from the native variety, and which (b) is subject to developmental change along the same lines. Second, the effects of and two possible explanations for this learner-specific variety of gestural anaphoric linkage will be discussed.

2.   An introduction to gestures and gestural cohesion

Gestures are defined in this paper as co-speech gestures or the (manual) movements speakers perform unwittingly while they speak as part of the expressive effort (cf. Kendon 1993; McNeill 1992). These movements are closely and systematically related to language and speech. Gestures are semantically co-expressive with speech, such that they often convey meaning also present in speech either iconically, or by way of spatial contiguity or indexicality. In (2), a speaker is performing a two-handed gesture which outlines a rectangular shape in space, iconically depicting the shape of the reception being talked about (see also Figure 2.2 in Section 5).

(2)   *la femme qui est [+ dans la réception +]*
      'the woman who is in the reception'

In addition to the overlap in content, the onset and duration of a gesture is tightly co-ordinated with the onset and duration of the referential expression in speech to which it is semantically linked. The square brackets in (2) mark the total extension of the manual movement, from preparation to retraction; boldface marks the stroke or the most forceful movement of the gesture; the underlined part marks a hold, i.e. a phase where the hands are held still in the air with maintained handshape (McNeill 1992); the plus (+) indicates a (non-measured) pause. As can be seen in (2), the hands move into position before *la réception* is uttered. The actual gestural movement occurs as the speaker is saying *dans la*, but the hands are maintained in the air until the noun *réception* has been uttered. Only then do the hands come down to rest. Without going into technical details, this tight semantic and temporal coordination suggests interdependence between the modalities.

Gestures and space also offer specific possibilities for cohesion. The mechanism of gestural cohesion rests on the continued or recurring gestural patterns – handedness, hand configuration, or specific spatial area – that manifest themselves over a stretch of discourse (labelled "catchments" by McNeill 2000) to reflect consistent visuospatial imagery running through a discourse segment. The close semantic and temporal co-ordination of gestures and speech also allow gestural patterns to be repeatedly associated with a given referential expression in speech (Levy & McNeill 1992; McNeill & Levy 1993). The association can be accomplished through iconicity or through indexicality such that a hand or a particular hand shape can come to represent a specific character throughout a story (Kendon 1972; McNeill & Levy 1993). Space may also be associated with particular referential expressions, such that one side of space is consistently indicated when speaking of *France* and the other when speaking of *Germany*, for instance. These different features are typically conflated such that spatial area and handedness coincide. Once

the association between a referent and a gesture-space complex is established, the activation of the spatial area or the hand (shape) results in the activation of the associated referent in discourse.

3.   Data background

The analyses in this paper are based on cross-sectional data of video recorded story retellings from five native speakers of Swedish learning French as a foreign language in a classroom setting, and five native speakers of French learning Swedish as a foreign language under similar circumstances. The subjects retold a cartoon story in both their native and their foreign language to a native speaker of the respective languages. Half of the subjects performed the task in L1-L2 order, the other half in L2-L1 order. The subjects thus acted as their own controls with respect to gesture production as well as with regard to story telling skills. The subjects were all at an intermediate level as established by a panel of native speakers, or at an early post-basic variety level (cf. Klein & Perdue 1997). This latter characterisation is based on the observation of some verbal and nominal morphology in the learner speech (typically finite forms of lexical verbs in the present tense, and some base form for past tense, tentative definiteness, gender, and number markings). This grouping notwithstanding, there was individual variation in the actual proficiency level. In addition to this data set, two learners from the Dutch and Swedish parts of the European Science Foundation project on adult second language acquisition have been analysed for gesture production.[1] These data serve as an important source of information and confirmation regarding a different learner type (untutored), a different proficiency level (lower), a different discourse type (personal narrative), and different source- and target languages (L1 Finnish > L2 Swedish, and L1 Moroccan Arabic > L2 Dutch). Note, however, that since there are no native baseline data for the gesture performance for the ESF data, they have not been included in the brief quantitative outline in Sections 4 and 5.

In the native and non-native data sets animate and inanimate referents were coded for referential status at the utterance level as introduced (new), maintained (old), re-introduced (new). *Introduced* refers to the first mention of an argument independently of its position in the clause (subject, object, or oblique argument). *Maintained* refers to the maintenance of the same referent in subject position in a subsequent clause. *Re-introduced* refers to the re-introduction of a previously mentioned referent subsequent to a clause containing a different subject.

Referents were also coded for whether they contributed to topic or focus elements. Topics were identified on the basis of the *aboutness* criterion: "The topic of a sentence is the thing which the proposition expressed by the sentence is ABOUT"

(Lambrecht 1994). Focus was defined in Lambrecht's terms as "that portion of a proposition which cannot be taken for granted at the time of speech. It is the UNPREDICTABLE or pragmatically NON-RECOVERABLE element in an utterance. The focus is what makes an utterance into an assertion" (1994:207). Note that topic-comment structure is not dependent on a referent's status as given-new, but rather reflects the underlying question that a text sets out to answer (von Stutterheim & Klein 1989).

All gestures in the native and non-native data were identified. Those gestures that co-occurred with referential expressions in speech were coded for location in gesture space. Note that gestures occurring with the predication of an utterance or the verb were not considered. Note also that in the L2 data, the same referential expression is often repeated with a simultaneous repetition of the gesture. Such speech-gesture clusters were only considered once.

4.   Gestural anaphoric linkage in L1

The general mechanism of gestural cohesion was briefly outlined above. The specifics of anaphoric linkage can be observed in fine-grained detail in the domain of referential movement in native production. When a referent is first introduced in discourse, speakers can accompany the NP with a pointing, deictic gesture, or with an iconic gesture depicting the referent. The gesture associates the referent with or localises it in the part of space indicated by the point or in the place where the iconic gesture was performed.[2] In Figure 1, a native speaker of Swedish localises both the *girl at the counter* and *the manager* at their first mention.

The girl is localised deictically, whereas *the manager* is anchored iconically such that a rectangular shape is outlined to symbolise his location. Speakers may point back to the locus in space previously established as associated with the referent when the referent is next mentioned. The anaphoric gesture re-activates the location and the referent associated with it (Levy & McNeill 1992; Marslen-Wilson, Levy, & Komisarjevsky Tyler 1982; McNeill & Levy 1993). The association between a referent and a locus serves not only to establish the referent, but rather to establish the larger scenes or spatial settings that the referent occupies. Indeed, when a localised referent moves, speakers often track it gesturally across gesture space, such that the end point of the trajectory becomes the referent's new anchor point. Notice, however, that such tracking gestures typically occur with the predication or the verbal element of an utterance and not with referring expressions (Levy & McNeill 1992). As these anchor points multiply, speakers charge concrete gesture space with referential meaning and turn it into a map of discourse (Liddell 1996; McNeill & Pedelty 1995). The relative location of a referent is usually irrelevant,
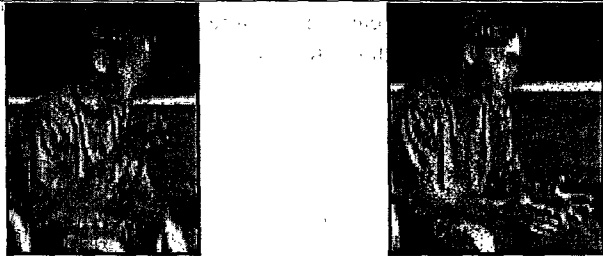
1 så visar hon de eh till eh då en [tjej] i disken	*2 + som tar detta då å tycker att de ser mycke mystisk ut + varefter hon tar de till [typ föreståndarn]*
1 'so she shows it uh to uh a [girl] at the counter	2 + who takes this then and thinks it looks very strange + whereupon she takes it to [sort of manager]'

* Only illustrated gestures are marked in the transcription

Figure 1. A native speaker of Swedish localises two referents at their introduction.* [] indicates the gesture phrase, **boldface** marks the gesture stroke, underlined marks gesture hold, and + indicates non-measured pause.

Table 1. Number of gestures occurring with introduced, maintained or re-introduced referential expressions in L1.

	Introduced	Maintained	Re-introduced
L1	43/171 (25%)	3/285 (1%)	18/128 (14%)

but once it has been established, it is constant and referential space takes on an absolute topological quality. This referential space or discourse map allows explicit and unequivocal visual co-reference.

A number of studies have shown that the gestural anchoring of referents is determined by their referential status as new or given, and on their status in the information structure. New (or re-introduced) referents which are part of focus elements tend to be localised, whilst given (maintained) referents which are part of topic elements tend not to be (Levy & McNeill 1992; Marslen-Wilson et al. 1982). These findings were replicated in this study (see Table 1). The speaker in Figure 1 gesturally anchors only the new referents; he does not refer back to them while they are maintained.

A similar effect for referential status has been observed for Sign Language, where concrete, specific referents with high thematic value are more likely to be localised than abstract or concrete referents with low thematic value (Engberg-

Pedersen 1993). This relationship between gestural cohesion and discourse status is further seen in the developmental effects of the deployment of anaphoric gestures. Only when children master the principles of discourse construction, information structure, and the use of appropriate grammatical means to organise these do they start producing gestures to mark referents in the adult fashion. In fact, they do not appear to produce adult-like gestural anaphora before the age of 12 (McNeill 1986).

5.   Gestural anaphoric linkage in L2

The results from studies of native discourse indicate that the deployment of anchoring and/or anaphoric gestures in L1 crucially depends on the very factors that cause learners problems in L2, viz. referential status and information organisation beyond the single utterance level. What does this mean for learners' use of anaphoric gestures in L2 production?

In Figure 2 a learner retells exactly the same episode in the story as in Figure 1, but this time in his foreign language.

The two new referents in this episode are treated as in L1: they are introduced with full NPs and gesturally anchored (*la femme* in Figure 2.1, and *le supervisé* in 2.6). In contrast to the native example, both these NPs are further specified by locative expressions which are also gesturally expressed and spatially anchored; (*la femme*) *dans la réception* in 2.2, and (*le supervisé*) *de la pharmacie* in 2.7. The referent *la femme* is subsequently maintained throughout this stretch by use of full NPs. With every such full NP, the learner points back to the anchor point associated with the referent. Notice that there are two female characters in this section. Since the first one is here labelled with a pronoun, *elle* in 2.1, the NP in Figure 2.3 could have been taken to re-introduce this first character. However, the anaphoric gestures occurring with the NPs consistently indicate the very location in space where the first NP was anchored, thereby indicating the co-referential reading. This example thus illustrates three points. First, learners over-mark referents gesturally in the sense that referents are first located at their introduction and then anaphorically indicated at every subsequent mention, even in an immediate context. Moreover, even at their introduction referents can be gesturally "over-marked". The specifying locative expressions (*dans la réception*) are sometimes also spatially anchored such that a single referent is gesturally marked twice. Second, referents that are over-marked in gesture are also over-marked in speech by nominal expressions. Maintained referents, for instance, are typically expressed with full NPs and also with anaphoric gestures. There is thus simultaneous over-marking in both modalities. Third, gestural anaphoric linkage allows unequivocal visual co-reference. Tables

1 eh elle a un une script + qui elle eh+ donné à la [+ femme+]	2 qui est [dans la réception+]	3 et [la femme] ne comprend pas +	4 et [la femme] +
1 'uh she has a a prescription + that she uh + give to the [+ woman+]	2 who is [in the reception +]	3 and [the woman] doesn't understand +	4 and [the woman] +

~ 5 [dans le réception]	6 eh donné la script eh de [+ le + supervisé]	7 + [de la pharmacie]
5 [in the reception]	6 uh give the prescription uh of [+ the + supervisor]	7 + [of the pharmacy]'

Figure 2. A learner of French localises referents gesturally in space at their introduction and at all subsequent mentions.

2–4 illustrate the first two points by summarising some differences between the patterns of gestural anaphoric linkage in L1 and L2 in the Swedish-French data.

While native speakers gesturally localise both newly introduced referents, and re-introduced referents, learners perform gestural anaphora with new, re-

Table 2. Number of gestures occurring with introduced, maintained or re-introduced referential expressions in L1 vs. L2.

	Introduced	Maintained	Re-introduced
L1	43/171 (25%)	3/285 (1%)	18/128 (14%)
L2	58/149 (39%)	31/208 (15%)	26/136 (26%)
L1 vs. L2	$\chi^2 = 3.2, p = 0.0736$	$\chi^2 = 28.85, p \leq 0.0001$	$\chi^2 = 0.59, p = 0.4424$

Table 3. Number of gestures occurring with topic or focus parts in L1 vs. L2.

	Topic	Focus
L1	10/64 (16%)	54/64 (84%)
L2	41/115 (36%)	74/115 (64%)
L1 vs. L2	$\chi^2 = 4.03, p = 0.0447$	$\chi^2 = 1.05, p = 0.3055$

Table 4. Number of gestures occurring with NPs or pronouns in L1 vs. L2.

	NP	Pron
L1	59/64 (92%)	5/64 (8%)
L2	108/115 (94%)	7/115 (6%)

introduced, *and* maintained referents. When the distribution of anaphoric gestures is considered across topic vs. focal parts of utterances, a similar pattern is seen. While the majority of the gestures occurring with referential expressions are part of the focus both in native and non-native discourse, learners perform a significantly greater number of gestures on topic material than native speakers. Finally, gestures occurring with referential expressions largely coincide with NPs both in native and non-native discourse. This means that in (nearly) all cases where anaphoric gestures occur, the indicated referents are expressed as NPs in speech. For native speech, this means with new or re-introduced referents in focus. For learner varieties, however, it means that in addition to contexts of new and re-introduced referents in focus, gestures occur with NPs expressing maintained referents in topic (Levy & McNeill 1992). In other words, just as learners' speech is characterised by nominal over-explicitness for maintained referents, so their gestural performance is over-marked with respect to the same expressions. Note that the same features can be observed in the two untutored learners in the ESF data set although they have not been included in the quantitative analyses. Despite the different discourse type (personal narrative), proficiency level (lower), and languages involved the tendency to gesturally refer to nominally maintained referents is the same.

An additional observation is that there is a clear tendency in the data for a re-duction of gestural anaphora with the presence of pronominal means (including zero anaphora) of co-reference. The decrease in localising gestures is not a matter of fluency or of lexical development. The learner in Figure 2 is relatively fluent in the sense that his hesitations are few and brief, and his use of the limited lexicon is quite creative. But the absence of pronouns is conspicuous, and his persistent use of NPs for maintained referents is still accompanied by localising gestures. In-versely, learners with relatively greater fluency problems but who *do* use anaphoric pronouns for maintained referents do not display anaphoric gestures with these expressions. The amount of localising-anaphoric gestures appears to depend cru-cially on the development of those grammatical means that allow a learner to mark referents as already established.

As a side remark, the tight link between gestures and NPs may at first sight seem surprising. In recognition of the powerful effect of visual co-reference, it is often assumed that gestures should function to disambiguate vague or underspec-ified expressions such as pronouns or zero anaphora.[3] In view of learners' poor command of pronouns and zero anaphora, one might therefore have assumed that where (erroneous) pronouns occur, they would be in need of gestural sup-port for anaphoric resolution. However, gestures do not in general co-occur with pronouns. The only time pronouns are accompanied by anaphoric gestures in na-tive or learner language is when the pronouns receive emphatic stress, or when the pronominal form itself is being negotiated, as in (3).

(3)   [*hans do*] *doktor eh* [*hennes*] *eh non hans hennes eh hans* [*doktor*]
      '[his do] doctor uh [her] uh no his her uh his [doctor]'

In this case, the gesture does not serve to identify a discourse referent. Instead, the gesture serves as a placeholder for or a localisation of *form*. The negotiation of the form is conducted at a meta-discursive level. The gesture is therefore not anaphoric at the narrative level, but does in fact serve to localise form at an even more abstract level of space.

To summarise thus far, the characteristic pattern for anaphoric linkage in the spoken learner variety is mirrored in gesture: maintained referents in topic are over-marked in speech by full NPs and in gesture by consistent anaphoric pointing. Put differently, the learner variety of gestural cohesion, or the gestural correlate to the spoken learner variety, differs systematically from L1 production in precisely the same way as the spoken variety. It is over-explicit, specifically with regard to maintained referents in topics. Moreover, it changes with the development of the grammatical means to achieve maintained reference. What, then, are the effects of the learner variety of gestural anaphoric linkage and what motivates it?

6.   Gestural anaphoric linkage – Is it communicatively motivated?

It has been argued that learners favour full NPs to avoid ambiguity in discourse caused by erroneous pronouns. Such an approach takes addressees and inter-pretability into account to explain learner behaviour. Williams (1988) has proposed that learners have two operating principles: a "hyper-clarity" principle ("be as ex-plicit as possible") for minimal ambiguity, and an economy principle for minimal redundancy. However, it might just as well be argued that learners cause hyper-ambiguity by using full NPs for maintained referents. By not demoting referents, learners violate a whole range of postulated principles for anaphora resolution: var-ious givenness hierarchies and accessibility scales (Ariel 1990; Chafe 1994; Givón 1984), the Gricean quantity principle (Grice 1975), the minimisation principle (Levinson 1998/1987, 2000), relevance (Sperber & Wilson 1986), principles of re-cipient design (Sacks & Schegloff 1979), and even the Zipfian principle of least effort (Zipf 1949). It poses obvious problems for interpretation when maintained referents are not formally distinguished from new ones, or when topic-focus status is only indicated by position (Klein & Perdue 1997). For instance, the repeated use of the NP *the woman* in Figure 2 is in fact an unsuccessful way of resolving co-reference since there are several female characters active in the context. Levinson suggests that the normal (read: L1) way of resolving co-reference is by applying the Minimisation principle which can be (carelessly) re-stated as 'say as little as possible, and interpret as widely as possible' ('Less is more'). Learners do the op-posite thing by applying a sort of Maximisation principle: 'say as much as possible, but interpret as narrowly as possible' ('More is less'). However, there is both psy-cholinguistic and interactional evidence suggesting that the outcome of the max-imisation principle, over-explicit referring expressions, are difficult to process for addressees. For instance, Cloitre and Bever (1988) showed that appropriately *light* referential expressions help identify referents better than inappropriately heavy or explicit ones. Pronouns referring to focal referents lead to faster retrieval of in-formation than explicit referring expressions. Similarly, there is ample evidence in the interactional data that native speakers experience problems processing and understanding the over-explicit learner narratives. The native speakers repeatedly engage in negotiations to clarify who actually did what in the story. The effect of over-explicitness in speech is thus ambiguity rather than clarity.

Gestural anaphoric linkage offers a potential solution to the communicative problem of hyper-ambiguous over-explicitness. In Figure 2, the deictic gestures al-low the addressee to identify the series of NPs *la femme* as being co-referential since the gestures indicate the same location in space for each expression. The anaphoric gestures allow full visual, unequivocal co-reference to be maintained throughout the sequence. In other words, only by being gesturally explicit is the learner achiev-ing hyper-clarity. Gestural anaphoric linkage can thus serve as a communicative
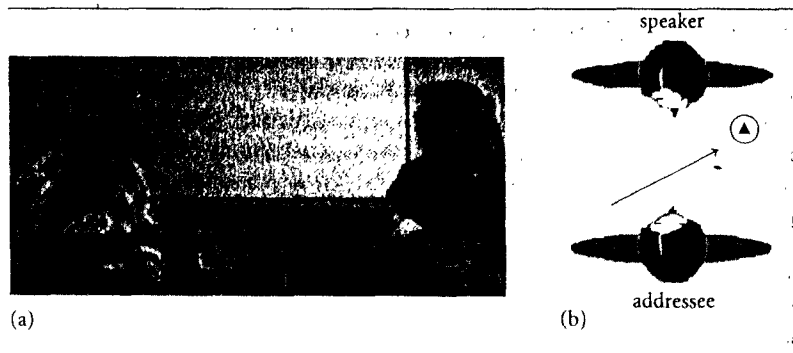
Figure 3. (a) The addressee (right) pointing to the same locus as indicated by the learner (left). (b) The map of discourse.

trade-off between over-explicit, ambiguous speech, and unambiguous gestures. It is compelling that deictic gestures with NPs often occur in sequences like these where there are several potential antecedents active (cf. Engberg-Pedersen 1993). It is thus tempting to assume that learners disambiguate gesturally what they cannot distinguish in speech.

This view of gestural cohesion rests on the assumption that addressees attend to gestures, specifically to (pointing) localising gestures and there is some evidence that they do. When cohesive gestures are deliberately mismatched with speech such that previously established locations are violated in re-activation, addressees find it hard to retell stories coherently (Cassell, McNeill, & McCullough 1999). Moreover, addressees themselves refer to the discourse maps established in interactional space, pointing back to loci established by the learner when negotiating a referent (Gullberg 1998). In doing so, they respect the absolute topological quality of the discourse map and the locations set up in the learners' space, and indicate exactly the same loci as indicated by the learners instead of setting up their own. The native speaker in Figure 3a points to the same location in space as the learner, thus referring to joint gesture and discourse space (Figure 3b).

This communicative, addressee-oriented account of bi-modal anaphora is intuitively attractive. However, it does not fully account for the link between NPs and gestures, nor for the fact that speakers, native and non-native alike, localise referents at their introduction. Moreover, it is not unequivocally clear that learners deliberately externalise and point to discourse for the benefit of the addressee in order to reduce ambiguity and allow for joint solutions (Gullberg 1998; Wilkes-Gibbs 1995). The problem here is well-known from the field of communication strategies: how intentional and/or aware does behaviour have to be in order to qualify as a

communication strategy (Gullberg 1998; Kasper & Kellerman 1997)? The localising gestures are clearly performed with a high degree of automaticity and little awareness. When asked if they remember performing these gestures, learners invariably say no – even when the loci and the referents associated with them have been the subject of extensive negotiation. As for other types of communication strategies, learners may 'subconsciously' perform gestures in contexts of potential ambiguity much in the same way as they can circumvent other communicative problems before these become manifest. In contrast to what is the case for many spoken strategies, the potentially intentional disambiguating, communicatively driven, effect of gestural anaphoric linkage can (and should) be tested empirically by manipulating addressees' visual access to learners' anaphoric gestures. Such a design will tell us to what extent learners actually consider their addressees when deploying anaphoric devices in speech and gesture in conjunction.

7.   Is it a reflection of learners' speech planning?

An alternative and more speculative view of anaphoric gestures considers the speaker's own perspective on over-marking in speech and gesture. In such an approach, gestures are not (solely) seen as interactive communicative solutions, but rather as a reflection of speakers' (and learners') cognitive efforts to construct utterances, and perhaps even global discourse structure beyond the single utterance level. Under this view, gestural anaphoric linkage could afford a new window into L2 speech planning and processing.

The field of gesture studies has recently seen the development of several theories assigning a role to gesture in speech production processes. McNeill's theory of gesture production, the so-called Growth Point theory (McNeill & Duncan 2000), is particularly interesting for the issue of discourse construction. The growth point in an utterance is similar to Vygotsky's 'psychological predicate', i.e. it represents a significant departure from what precedes in an immediate context, and represents the new "idea". The growth point is materialised in gesture and speech simultaneously – in linear analytic form in speech, and in global synthetic form in gestures, each modality contributing that part of the information for which it is best suited. Since gestures reflect a new idea, they occur either with predications or with referential expressions in focus, as seen above (Levy & McNeill 1992). This account indirectly explains why there are no gestures with pronouns. Since pronouns inherently express presupposed material that is not part of the growth point, they should not (and do not) receive gestural expression.

Another potential consequence of this view is that the units in speech that co-occur with gestures reflect units planned for speech execution. I.e. if a gesture re-

flects a growth point or new material to be expressed, then it might also be seen as a reflection of a planning unit. What does this mean for early L2 acquisition? Learner speech is typically non-fluent and proceeds in small units interrupted by frequent pauses, sometimes between every unit in an utterance. Similarly, several observers have noted that learners at such stages of acquisition typically produce gestures on almost every argument of an utterance (Gullberg 1998; Nobe 1993). The learner in (4) produces a gesture on the predication, another one on the referential expression in direct object position, as well as one on the oblique object.

(4)   *hon eh + [ger] hon ger eh [pa papper] papper eh [till apotekarin]* –
        she uh + [gives] she gives uh [pa paper] paper uh [to the pharmacist]

If gestures mark growth points or idea units, then more idea units are needed to express all events at early stages of discourse production in L2 as in (4). Gestures would then be a reflection of learners' L2 speech planning proceeding by smaller units (Nobe 1993). With increasing fluency and automaticity in speech production (lexical retrieval, grammatical encoding), learners' gestures package more meaning units into one single gesture (e.g. the action of giving and the paper in one gesture). Returning to referential movement, a gesture on a maintained referent expressed as an NP in topic could then be taken to mean that this entity has been planned as a separate unit, much as if it were a new idea. This line of reasoning opens up a range of new questions pertaining to how a processing approach could account for the use of full NPs in speech, and the late acquisition of pronouns; to whether and how gestures and externalisation of discourse help to reduce learners' cognitive load. These issues remain to be explored both in the field of gesture studies and in SLA research.

8.   Final remarks

Both communicative and cognitive mechanisms are likely to play an important role in the production of over-explicit anaphoric linkage in speech and gesture. Although much remains to be done in this area, especially with respect to the cognitive, speaker-internal perspective, it is hopefully clear by now that gesture analysis can offer important insights into both aspects of anaphoric linkage. In addition, gestural cohesion may also shed new light on stabilisation vs. learners' continued development. Anaphoric gestures offer a tempting explanation for why learners 'can' stabilise at levels where their speech alone is not communicatively efficient. As shown in this paper, the combination of over-explicit anaphora in both modalities provides learners with greater communicative efficiency then they would have in speech alone. The assumption that communicative efficiency prompts stabili-

sation is not new. The classical description of the communicatively efficient but formally unsuccessful learner Wes (Schmidt 1983) certainly reinforced this conclusion. However, it is not a foregone conclusion that communicative efficiency is harmful to continued formal acquisition. Communicative efficiency may also be seen as an important contributing factor in the development of language, namely in the sense that it improves learners' possibilities to produce "comprehensible output" (Swain 1985). Comprehensible output has been suggested as one of the driving forces in language development. If the view is that you have to talk to learn how to talk (Bruner 1990), then anaphoric gestures certainly help you do this in interaction, whether you see them as communicatively or processing driven. In this sense they may actually contribute to the formal development of linguistic means for anaphoric linkage.

Notes

1. These data were collected within the European Science Foundation project *Second Language Acquisition by Adult Immigrants* by the teams of Universiteit Brabant, Tilburg, the Netherlands, and Göteborgs Universitet, Göteborg, Sweden. The data were kindly made available to me by the Central Data Archives of this project, located at the Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands, and by the Dept. of Linguistics, Göteborgs Universitet, Sweden.

2. The term localisation is borrowed from Sign Language where a similar phenomenon is part of the linguistic, grammatical system. Nominal signs are localised or associated with spatial loci at their introduction. Pronominal signs directed towards these loci are interpreted as co-referential with the nominal. Signs for verbal components are also oriented with respect to these anchor points such that localisation is part of the grammatical means to achieve agreement (Engberg-Pedersen 1993; Liddell 1996).

3. See for instance Levinson (1998/1987:598) on the extensive usage of zero anaphora in Guugu Yimithirr: "[...] it is hardly surprising that a language with zero anaphora and no verbal agreement would find an ancillary channel of gestural information very useful indeed [...]." It is important to note, however, that this disambiguation is not a function of gestures co-occurring with zero anaphora. Gestures that can disambiguate zero anaphora are instead iconic gestures that appear with the predication of an utterance, rather than with the referring expression. Iconic gestures can indirectly express agent or patient by being oriented towards previously established locations in space much as in Sign Language. And indeed, in the GY data pointing gestures co-occur only with NPs, demonstratives (±N), or location adverbials, similarly to the learner data under consideration, and not with zero anaphora or indeed with pronouns themselves. Note also that the disambiguating function of gestures with respect to expressions like demonstratives is not under discussion here, since this is an exophoric use of gestures, not an endo- or anaphoric one.

References

Ahrenholz, B. (1998). *Modalität und Diskurs: Instruktionen auf deutsch und italienisch. Eine Untersuchung zum Zweitspracherwerb und zur Textlinguistik*. Tübingen: Stauffenburg.

Ariel, M. (1990). *Accessing noun phrase antecedents*. New York: Routledge.

Bruner, J. (1990). *Acts of meaning*. Cambridge, MA: Harvard University Press.

Carroll, M., Murcia-Serra, J., Watorek, M., & Bendiscoli, A. (2000). The relevance of information organization to second language acquisition studies: The descriptive discourse of advanced adult learners of German. *Studies in Second Language Acquisition, 22*(3), 441–466.

Cassell, J., McNeill, D., & McCullough, K.-E. (1999). Speech-gesture mismatches: Evidence for one underlying representation of linguistic and nonlinguistic information. *Pragmatics & Cognition, 7*(1), 1–33.

Chafe, W. (1994). *Discourse, consciousness, and time*. Chicago: University of Chicago Press.

Cloitre, M. & Bever, T. G. (1988). Linguistic anaphors, levels of representation and discourse. *Language and Cognitive Processes, 3*(4), 293–322.

Engberg-Pedersen, E. (1993). *Space in Danish Sign Language*. Hamburg: Signum.

Extra, G., Strömqvist, S., & Broeder, P. (1988). Pronominal reference to persons in adult second language acquisition. In P. Broeder, G. Extra, R. v. Hout, S. Strömqvist, & K. Voionmaa (Eds.), *Processes in the developing lexicon* (pp. 86–113). Strasbourg, Tilburg, Göteborg: European Science Foundation.

Fox, B. (1987). *Discourse structure and anaphora*. Cambridge: Cambridge University Press.

Givón, T. (1984). Universals of discourse structure and second language acquisition. In W. E. Rutherford (Ed.), *Language universals and second language acquisition* (pp. 109–136). Amsterdam: Benjamins.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Speech acts* (pp. 41–58). New York: Academic Press.

Gullberg, M. (1998). *Gesture as a communication strategy in second language discourse. A study of learners of French and Swedish*. Lund: Lund University Press.

Hendriks, H. (2000). The acquisition of topic in Chinese L1 and French L1 and L2. *Studies in Second Language Acquisition, 22*(3), 369–397.

Jin, H. G. (1994). Topic-prominence and subject prominence in L2 acquisition: Evidence of English-to-Chinese typological transfer. *Language Learning, 44*, 101–122.

Kasper, G. & Kellerman, E. (1997). Introduction: approaches to communication strategies. In G. Kasper & E. Kellerman (Eds.), *Communication strategies: Psycholinguistic and sociolinguistic perspectives* (pp. 1–13). London: Longman.

Kendon, A. (1972). Some relationships between body motion and speech: An analysis of an example. In A. W. Siegman & B. Pope (Eds.), *Studies in dyadic communication* (pp. 177–210). New York: Pergamon.

Kendon, A. (1993). Human gesture. In K. R. Gibson & T. Ingold (Eds.), *Tools, language and cognition in human evolution* (pp. 43–62). Cambridge: Cambridge University Press.

Klein, W. & Perdue, C. (1997). The basic variety (or: Couldn't natural languages be much simpler?). *Second Language Research, 13*(4), 301–347.

Lambrecht, K. (1994). *Information structure and sentence form. Topic, focus, and the mental representation of discourse referents*. Cambridge: Cambridge University Press.

Levinson, S. C. (1998/1987). Minimization and conversational inference. In A. Kasher (Ed.), *Presupposition, implicature and indirect speech acts* (pp. 545–614). London: Routledge.

Levinson, S. C. (2000). *Presumptive meanings. The theory of generalized conversational implicature*. Cambridge, MA: MIT Press.

Levy, E. T. & McNeill, D. (1992). Speech, gesture, and discourse. *Discourse Processes, 15*(3), 277–301.

Liddell, S. K. (1996). Spatial representations in discourse: comparing spoken and signed language. *Lingua, 98*(1–3), 145–167.

Marslen-Wilson, W., Levy, E., & Komisarjevsky, Tyler, L. (1982). Producing interpretable discourse: The establishment and maintenance of reference. In R. J. Jarvella & W. Klein (Eds.), *Language, place, and action. Studies in deixis and related topics* (pp. 339–378). Chichester: John Wiley.

McNeill, D. (1986). Iconic gestures in children and adults. *Semiotica, 62*(1/2), 107–128.

McNeill, D. (1992). *Hand and mind. What the hands reveal about thought*. Chicago: Chicago University Press.

McNeill, D. (2000). Catchments and contexts: nonmodular factors in speech and gesture production. In D. McNeill (Ed.), *Language and gesture* (pp. 312–238). Cambridge: Cambridge University Press.

McNeill, D. & Duncan, S. D. (2000). Growth points in thinking-for-speaking. In D. McNeill (Ed.), *Language and gesture* (pp. 141–161). Cambridge: Cambridge University Press.

McNeill, D. & Levy, E. T. (1993). Cohesion and gesture. *Discourse Processes, 16*(4), 363–386.

McNeill, D. & Pedelty, L. L. (1995). Right brain and gesture. In K. Emmorey & J. S. Reilly (Eds.), *Language, gesture, and space* (pp. 63–85). Hillsdale, NJ: Erlbaum.

Muñoz, C. (1995). Markedness and the acquisition of referential forms. The case of zero anaphora. *Studies in Second Language Acquisition, 17*(4), 517–527.

Muñoz, C. (2000). *The over-explication of personal reference in L2 narratives*. Paper presented at the Euresco Conference 'The structure of learner language', San Feliu de Guixols, October 7–12, 2000.

Nobe, S. (1993). Cognitive processes of speaking and gesturing: A comparison between first language speakers and foreign language speakers. Unpublished Masters thesis, Dept. of Psychology, University of Chicago.

Pekarek, S. (1999). Linguistic forms and social interaction: Why do we specify referents more than is necessary for their identification? In J. Verschueren (Ed.), *Pragmatics in 1998. Selected papers from the 6th International Pragmatics Conference* (pp. 427–448). Antwerpen: International Pragmatics Association.

Sacks, H. & Schegloff, E. A. (1979). Two preferences in the organization of reference to persons in conversation and their interaction. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology* (pp. 15–21). New York: Irvington.

Schmidt, R. W. (1983). Interaction, acculturation, and the acquisition of communicative competence: A case study of an adult. In N. Wolfson & E. Judd (Eds.), *Sociolinguistics and language acquisition* (pp. 137–174). Rowley, MA: Newbury House.

Sperber, D. & Wilson, D. (1986). *Relevance. Communication and cognition*. Oxford: Blackwell.

Strömqvist, S. & Day, D. (1993). On the development of narrative structure in child L1 and adult L2 acquisition. *Applied Psycholinguistics, 13*(2), 135–158.

Swain, M. (1985). Communicative competence. Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Rowley, MA: Newbury House.

von Stutterheim, C. & Klein, W. (1989). Referential movement in descriptive and narrative discourse. In R. Dietrich & C. F. Grauman (Eds.), *Language processing in social context* (pp. 39–76). Amsterdam: North-Holland.

Wilkes-Gibbs, D. (1995). Coherence in collaboration: Some examples from conversation. In M. A. Gernsbacher & T. Givón (Eds.), *Coherence in spontaneous text* (pp. 239–267). Amsterdam: Benjamins.

Williams, J. (1988). Zero anaphora in second language acquisition. *Studies in Second Language Acquisition, 10,* 339–370.

Zipf, G. K. (1949). *Human behavior and the principle of least effort: An introduction to human ecology.* New York: Hafner.