The role of iconicity and simultaneity for efficient communication: The case of Italian Sign Language (LIS)

Anita Slonimska [a,b,*], Asli Özyürek [b,c], Olga Capirci [a]

[a] *Institute of Cognitive Sciences and Technologies (ISTC), National Research Council (CNR) of Italy, Via S. Martino della Battaglia, 44, 00185 Rome, RM, Italy*
[b] *Radboud University, Centre for Language Studies, Erasmusplein 1, 6525 HT Nijmegen, the Netherlands*
[c] *Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, the Netherlands*

ABSTRACT

A fundamental assumption about language is that, regardless of language modality, it faces the *linearization problem*, i.e., an event that occurs simultaneously in the world has to be split in language to be organized on a temporal scale. However, the visual modality of signed languages allows its users not only to express meaning in a linear manner but also to use iconicity and multiple articulators together to encode information simultaneously. Accordingly, in cases when it is necessary to encode informatively rich events, signers can take advantage of simultaneous encoding in order to represent information about different referents and their actions simultaneously. This in turn would lead to more iconic and direct representation. Up to now, there has been no experimental study focusing on simultaneous encoding of information in signed languages and its possible advantage for efficient communication. In the present study, we assessed how many information units can be encoded simultaneously in Italian Sign Language (LIS) and whether the amount of simultaneously encoded information varies based on the amount of information that is required to be expressed. Twenty-three deaf adults participated in a *director-matcher* game in which they described 30 images of events that varied in amount of information they contained. Results revealed that as the information that had to be encoded increased, signers also increased use of multiple articulators to encode different information (i.e., kinematic simultaneity) and density of simultaneously encoded information in their production. Present findings show how the fundamental properties of signed languages, i.e., iconicity and simultaneity, are used for the purpose of efficient information encoding in Italian Sign Language (LIS).

1. Introduction

In order to share a thought with others through language we need to decompose our message into smaller information units and then organize these units on a linear scale. While all languages face this *linearization problem* (Levelt, 1981), the impact of it might be different based on the main modality the language is realized in. "For spoken language, linearization is an absolute requirement" (Levelt, 1980, p.153) given that a thought has to be split in a strictly sequential manner (i.e., sounds arranged in words, words - in sentences, sentences - in discourse etc.).[1] Although Levelt (1980, p.156) speculates that linearization problem might also stand for signed languages as it is for spoken languages,

empirical research on this subject and to what extent not only linearity but also simultaneity is fundamental for sign languages is missing. In sign languages, affordances of the visual modality allow packaging multiple units of information not only sequentially, but also simultaneously, through the use of multiple articulators. Simultaneity of the articulators also allows messages to be expressed in an iconic manner (i.e., *diagrammatic iconicity*, Perniss, 2007a, 2007b; Risler, 2007; Taub, 2001).

Following cognitive linguistics approach language is tightly interrelated with general cognitive processes of human mind (Croft & Cruse, 2004; Elman et al., 1996). The linguistic signal is fast and fleeting in both the acoustic and visual modalities. This implies that information

---

* Corresponding author at: Institute of Cognitive Sciences and Technologies (ISTC), National Research Council (CNR) of Italy, Via S. Martino della Battaglia, 44, 00185 Rome RM, Italy.
E-mail addresses: anita.slonimska@istc.cnr.it (A. Slonimska), Asli.Ozyurek@mpi.nl (A. Özyürek), olga.capirci@istc.cnr.it (O. Capirci).

[1] Note that some simultaneity in spoken communication is possible when taking into account that information can be conveyed in gesture and/or prosody alongside speech (Clark, 2016; Kendon, 2001, 2014; McNeill, 2008; Wilson & Wharton, 2006).

has to be encoded efficiently not only regarding the effort of the producer but also in relation to the informativity for the perceiver to understand it (Grice, 1975). Thus, encoding information both in acoustic and visual modalities face the efficiency problem. To what extent using simultaneity of the articulators to express information simultaneously and iconically is recruited for efficient information encoding in sign languages is underexplored.

In the present paper we are interested in how the capacity of sign languages to encode multiple information units simultaneously is modulated when producing informative messages of different informative density in Italian Sign Language (LIS). We hypothesize that in signed languages efficient communication about events perceived simultaneously in the world can be achieved by strategic use of simultaneous encoding of information. We predict that signers will make more use of simultaneous information encoding as a function of the amount of information that is required to be encoded for efficient communication.

In the following paragraphs, we will describe first how efficient information encoding has been addressed in spoken languages and then proceed to elaborate the notions of iconicity and simultaneity in sign languages. We will then test whether signers take advantage of iconic and simultaneous encoding of information when dealing with highly informative messages in a semi-naturalistic experiment.

1.1. Information packaging for efficient communication

The pressure to produce and process information rapidly has far-reaching implications on how languages are organized in regard to their structure (Christiansen & Chater, 2016; Hawkins, 2004; Jaeger and Tily, 2011; Lu, Xu, & Liu, 2016). In spoken languages efficient communication is achieved by reducing dependency length distances (Gibson et al., 2019; Hawkins, 2004) and reducing linguistic forms when possible (Aylett & Turk, 2004; Jaeger, 2006, 2010; Jaeger and Tily, 2011; Jaeger & Levy, 2007; Piantadosi, Tily, & Gibson, 2011; Mahowald, Fedorenko, Piantadosi, & Gibson, 2013). Considering that language encoding and decoding is dependent on cognitive constraints common to all language users regardless of the linguistic modality (Christiansen & Chater, 2016), it is plausible to assume that similar strategy like reducing dependency distances would be also used in signed languages for achieving efficient communication. On the other hand, efficiency in regard to reduction of linguistic forms in spoken languages has been mainly concerned with reduction and omission of function words (Ferreira & Dell, 2000; Jaeger & Levy, 2007; Jaeger and Tily, 2011; Race & MacDonald, 2003; Tagliamonte, Smith, & Lawrence, 2005) which in sign languages mostly do not exist (Fang, Gao, & Zhao, 2006). Thus, this strategy is less likely for sign languages. For this reason, in this paper we focus on dependency distance reduction only.

It has now been established that users of spoken languages tend to cluster words that are syntactically and semantically related (Bybee, 1985, 2013; Futrell, Mahowald, & Gibson, 2015; Gibson, 2000, Gibson 1998; Gildea & Temperley, 2010; Grodner & Gibson, 2005; Hawkins, 2004; Liu, Xu, & Liang, 2017; Temperley, 2007; Temperley & Gildea, 2018). It is argued that this strategy reduces cognitive load on working memory when forming relations among adjacent words in both production and comprehension (Ferrer i Cancho, 2004; Liu, 2008, Ferrer i Cancho & Liu, 2014; Gibson et al., 2019; Hawkins, 2004, 2001; Jiang & Liu, 2015). By means of dependency distance reduction, efficient communication can be achieved given that access to syntactic and semantic representation is provided as fast as possible (Hawkins, 2004, p.9) both for producers and comprehenders.

It would be expected that encoding semantically related information closer together also in signed languages would lead to reduction of cognitive load as in spoken languages. Given that sign languages allow encoding related information not only sequentially but also

simultaneously, it is plausible to assume that simultaneous encoding of information could be exploited for achieving efficient communication. Accordingly, the differences in affordances of language modality could potentially lead to different strategies of how information is packaged for efficient communication.

1.2. Iconicity and simultaneity in sign languages

The role of iconicity i.e., resemblance between the meaning and its linguistic form (Taub, 2001; Wilcox, 2004) has become a hot topic in language research in recent years and has been finally brought into prominence as a general property of language (Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015; Perniss, Thompson, & Vigliocco, 2010; Perniss & Vigliocco, 2014). Iconicity plays a fundamental role in language from the very beginning of its development as it provides motivated links to the experience in the world (Ortega, Sümer, & Özyürek, 2017; Ortega, 2017; Perniss & Vigliocco, 2014). Iconicity can be expressed at the lexical level, i.e., encoding correspondence between visually perceptual features of the sign and its referent (e.g., sign for *a bird* resembles a beak of the bird in Italian Sign Language), also called as *imagistic iconicity* (Russo, 2004; Taub, 2001). Similarly, also in spoken languages we can find words that perceptually resemble their meaning like onomatopoeias *meow* and *bang* in English or ideophones *sinisinisini* "closely woven", *saaa* "cool sensation" in Siwu (Dingemanse, 2011; Dingemanse et al., 2015).

Another type of iconicity found in both signed and spoken languages is called *diagrammatic iconicity* (Haiman, 1985; Perniss, 2007a, 2007b; Pietrandrea & Russo, 2007; Taub, 2001; Ungerer, 1999). This type of iconicity is not linked to the perceptual but to structural resemblance between form and meaning. Namely, the relationship between parts of a specific meaning motivates the relationship between parts of its linguistic form. For example, structural resemblance can be identified in previously mentioned sign *bird*. In LIS it is performed by closed fist and extended thumb and index finger that open and close. This sign is positioned in front of the signer's mouth with fingers facing outwards. Here we see that even the smallest building blocks of the sign, i.e., parameters: handshape, location, orientation & movement are not meaningless units but instead have imagistic iconic properties (Boyes Braem, 1980; Emmorey, 2014; Occhino, 2017) while their relation to each other is diagrammatically iconic (Emmorey, 2014; Lepic & Occhino, 2018; Taub, 2001). Furthermore, diagrammatic iconicity is also present in compounding when two signs are combined or merged for a new meaning, e.g., in ASL sign for *inform* consists of two signs *know* and *offer* articulated in a single smooth movement (see Lepic, 2015 for this and other forms of compounding in ASL). In spoken languages, too, we can observe diagrammatic iconicity in compounding (Lepic, 2015; Ungerer, 1999)., e.g., in English, word *glamping* consists of two words, *glam(orous)* and *camping,* which have to be interpreted in relation to each other in order to derive the meaning of the compounded word (Lepic, 2015, p.212).

Diagrammatic iconicity goes beyond lexical level and allows establishing meaningful semantic and syntactic relations. In spoken languages diagrammatic iconicity has been acknowledged as integral part of grammatical structure (Haiman, 1985). However, it is mainly limited to iconicity of sequence as in "the order of statements in a narrative description corresponds to the order of the events they describe" (Haiman, 1980, p.516). In spoken languages diagrammatic iconicity is strictly linear considering that "since it is impossible to say everything at once, words must appear in a certain order" (Haiman, 1980, p.528). When it comes to signed languages, the strictly linear nature of information unfolding can be overcome. Given their visual modality and accordingly the use of not only linear but also three dimensional space for linguistic encoding diagrammatic iconicity can be taken advantage

a.

b.
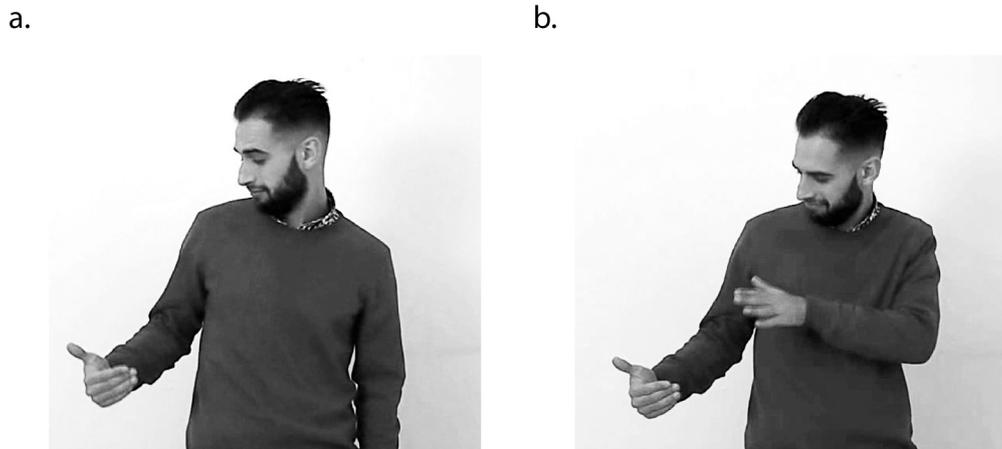
Fig. 1. a) A signer encoding a woman who is holding a dog. Woman (agent) is mapped onto the body of the signer marked by torso, head, eye gaze, and face expression. The holding action is encoded on the right hand.
b) A signer superimposing an additional action of caressing (left hand) onto previously encoded representation of the woman holding the dog.

of in order to sign multiple things simultaneously,[2] while establishing motivated relations between them (Risler, 2007). For example, events that are perceived simultaneously in the world e.g., events involving multiple referents and/or their actions and interaction can be encoded in such way that perceptual simultaneity consisting of multiple meaning elements (i.e., information units) is encoded simultaneously in linguistic form as well. This can manifest in different ways depending on the type and complexity of the event.

For example, in order to encode an agent, his/her action/s and a patient, e.g., a woman holding a dog, a signer after introducing both referents with lexical signs (*woman, dog*) could then take on the role of the woman (agent) by mapping the woman onto the signer's body (e.g., through face expression, eye gaze, and/or torso) and at the same time encode the holding action by one of the hands representing holding (see Fig. 1a). Note that in this example there are two schematized elements of the same event encoded, i.e., the agent and the action in respect to the agent. The dog is only available implicitly while the woman and the holding action are encoded explicitly. If the event to be communicated consisted of a woman holding AND caressing the dog, the signer could simply superimpose another information unit (i.e., *caressing*) by doing the caressing action with the other hand in respect to the location of the referent it should be acting on, in this example - a dog being held by the woman (see Fig. 1b). A signer mapping the referent onto his/her body and producing actions in relation to this embodied representation is known as *constructed action* (CA) (Cormier, Smith, & Sevcikova-Sehyr, 2015; Metzger, 1995). Although not acknowledged in previous literature such constructions involve not only imagistic properties of individual lexical items but diagrammatic iconicity as well. Namely, in such constructions each articulator can be used to encode different information about a referent simultaneously while being in a meaningful relation to each other through iconic schematization as exemplified above.

Alternatively, it would also be possible to use constructed action to encode information units consecutively, e.g., first encoding holding action only and then encoding caressing action only in location where the signer had just encoded holding action. In this scenario, diagrammatic iconicity would be maintained, considering that both actions are linked with motivated use of space (they are performed based on the location and properties of the patient). However, simultaneity would be lost and accordingly a single representation would be split in two parts (first holding, then caressing). These two parts would be nevertheless iconically linked between each other. Finally, encoding could be also done using specific lexical signs (i.e., *to hold, to caress*) articulated in neutral space in front of the signer and thus without resorting to diagrammatic iconicity at all. In this scenario, each information unit would be presented separately from others without establishing iconic link between the signs. As a result, such encoding would not form nor single nor split iconic representation from explicitly encoded information as in previous two cases. Such encoding would rely on maintaining previously presented information in working memory only and interpreted exclusively based on specific linguistic structure as would be done in spoken languages.

A more complex event could involve two referents' actions interacting with each other. In this scenario, mapping one of the referents and its actions onto the signer's body would encode only part of the event. The action/s of the other referent would then need to be encoded in a consecutive manner. It is then possible for signers to use a set of articulators to encode one referent and/or his actions, while other articulators can be used to encode another referent/s and/or actions. Such encoding can be achieved through *body partitioning* (Dudis, 2004). Dudis (2004) discusses partitionable zones of signer's body that allow the production of distinct information elements (i.e., information units) directly accessible to interlocutors. As main partitionable zones Dudis considers both hands and face. Face can be divided further into 2 zones that can be used to express different information units – face expression (including eye gaze) and oral articulators (e.g., mouth). Given that signers can partition their body to create distinct information units it is possible to encode information involving multiple referents simultaneously. Let's take an example from Pizzuto, Rossini, Sallandre, and Wilkinson (2006) of a signer encoding the boy holding a dog and the dog licking the cheek of the boy in LIS (see Fig. 2). In this example there are 4 main pieces of information (*distinct elements* in terms of Dudis, 2004) explicitly encoded, namely 1) *the boy* who 2) *is holding* 3) *a dog* and the dog who 4) *is licking* the boy. The face of the signer is partitioned in such way that the upper part of the face including eye gaze is the boy gazing at the dog while the face expression including the mouth refers to the dog (thus, we have evidence that eyes can be partitioned from the rest of the face as well). Note that the head is to be considered that of the boy because the sign for *licking* is produced on the cheek of the signer, i.e., the cheek of the boy who is looking at the dog. The lexical sign for *licking*

Fig. 2. A signer encoding two simultaneously occurring events; a boy holding a dog and the dog licking the boy. The boy is mapped onto the body of the signer and marked by torso and eye gaze. Boy's holding action is encoded on the left hand. The dog is marked by face expression of the signer and mapping between the dog's mouth and the signer's mouth. Licking action is encoded on the right hand and mouth of the signer. Figure has been recreated for demonstration purposes. See original in Pizzuto et al. (2006, p.483).

(right hand sign) can be diagrammatically positioned in respect to the constructed action of the boy holding the dog. Here we can observe how iconicity affords integration of multiple elements that can be combined together to create single representation of the event. If we take into account the eye gaze and location of licking as crucial information that has to be encoded then we have 6 information units (the boy, holding, the dog, licking, the cheek, and eye gaze direction of the boy) available at the same time.

To summarize, in sign languages different articulators encoding different information are all interpreted in relation to each other in given moments in time. The interpretation of such simultaneous constructions is possible due to diagrammatic iconicity that establishes these relations and links them to semantic representation (Risler, 2007, p.75). In the next section we briefly review previous research on how amount of simultaneously encoded information has been studied.

1.2.1. Simultaneous information encoding in sign languages

In research on simultaneity we can encounter different kinds of terminology that refers to different phenomena. For this paper, it is important to differentiate between *encoding of simultaneity* and *simultaneous encoding*. While the former refers to the grammatical structure of encoding referents and/or actions that appear/happen at the same time by using various linguistic strategies (in both spoken and signed languages) the latter refers specifically to the property of signed languages to encode different semantically or discourse-related information simultaneously. The present paper focusses on the latter.

Due to the affordances of the modality spoken languages encode simultaneity in a linear manner, i.e., use specific lexical items (e.g., while, and, as) to link pieces of meaning together to transmit different information that is semantically simultaneous (Morgan, 2002). The same applies if speakers want to elaborate further on the same event. In spoken languages, this has to be done sequentially. Also in signed languages it is possible to use lexical markers or sequential grammatical structures like "sandwiching", which is the doubling of the verb, to indicate simultaneity of the events (Fischer & Janis, 1990; Morgan et al., 2002; Napoli & Sutton-Spence, 2014). Alternatively, due to use of multiple articulators and signing space simultaneous encoding can be employed (Napoli & Sutton-Spence, 2010, Morgan et al., 2002). In the present paper we are concerned with simultaneous encoding of information related to the events that are perceived simultaneously.

Note that simultaneous encoding in signed languages is possible not only for events involving multiple referents engaged in simultaneous activities but also for the information of the same referent like encoding different perspectives of the same referent/action (Perniss, 2007a, 2007b). Simultaneous encoding can be also used to mark other properties of the discourse (e.g., various type of listing (Liddell, Vogt-Svendsen, & Bergman, 2007)). Vermeerbergen, Leeson, and Crasborn (2007) devote an entire volume to simultaneity, the first and currently only volume devoted to this topic. They differentiate between specific combinations of articulators used for specific purposes: manual simultaneity where each hand encodes different information; manual-oral simultaneity where mouthing's and mouth gestures contribute to the meaning expressed by the hands; simultaneous use of manual and-non manual articulators to encode different perspectives of the same event or multiple events simultaneously. In the present paper, we are not concerned with grammatical structure of encoding of events that are perceived simultaneously, but rather with the amount of information that can be packaged simultaneously to encode such events.

There are some descriptive studies that assess the capacity and/or limits of amount of information that can be encoded simultaneously in a sign language (Dudis, 2004; Napoli & Sutton-Spence, 2010; Perniss, 2007a, 2007b; Pizzuto et al., 2008; Risler, 2007; Sutton-Spence & Braem, 2013). Note that all research on simultaneous information encoding has been exclusively based on narrative data and has been mostly descriptive with little to no consideration of the functional role of iconicity in linking different information units into single representation. One hypothesis, however, has been proposed in the past. Namely, that simultaneity is used for communicative efficiency and informativeness. Perniss (2007a, 2007b) explores simultaneous encoding of the same event from different perspectives, observer and character perspective (i.e., *non-prototypical alignment*), and proposes that such simultaneity is used to achieve communicative efficiency and informativeness for coherence and ambiguity reduction in discourse. However, also this research in based on narrative data. This factor may create an issue in assessing the functional role of simultaneous information encoding in signing. Namely, narrating a story requires a signer to be not only informative (i.e., referential function) but also be expressive and entertaining (i.e., evaluative function) in their production (Labov & Waletzky, 1967; Özyürek & Trabasso, 1997). It has been argued that highly iconic strategy like constructed action, which is directly linked to simultaneous information encoding as described in previous paragraphs and to *non-prototypical alignment*, is used to fulfil the evaluative function in narratives as opposed to referential function (see Rogers, 2012 for a review). Accordingly, it is impossible to tease apart whether simultaneity is used as a narrative device to enhance the evaluative function of narration or whether it is used to increase informative efficiency.

In light of communicative efficiency, description of simultaneity is mostly concerned with encoding the same information from different perspectives as described above (Perniss, 2007a, 2007b) or alternatively with simultaneity of multiple predicates (Napoli & Sutton-Spence, 2010; Risler, 2007). Accordingly, there is no consideration that a single predicate might contain multiple semantically distinct information units which are encoded simultaneously and linked through diagrammatic iconicity.

Napoli and Sutton-Spence (2010) assess limitations on simultaneity in American Sign Language (ASL) by counting how many "propositions"[3] can be encoded simultaneously. Authors note that simultaneous encoding of two propositions is quite common in everyday signing while simultaneous production of three or four propositions is found in creative signing like poems. Authors find maximum of four propositions that

---

[3] Napoli and Sutton-Spence (2010) define *proposition* as follows: "[..] a proposition is a predicate and its constellation of arguments (Johnston, Vermeerbergen, Schembri, & Leeson, 2007) and is free of internal conjunction." (p.650)

can be encoded simultaneously. Although, Risler (2007) does not count the number of information units that is possible to encode simultaneously, similarly like Napoli and Sutton-Spence (2010) she accounts for predicative simultaneity and describes how two processes[4] can be encoded simultaneously in French Sign Language (LSF). As a result, the full range of simultaneously encoded information is not accounted for. For instance, in the example described in Fig. 2, Napoli and Sutton-Spence (2010) would only account for two simultaneous "propositions", and Risler (2007) would account for two "processes", disregarding the richness of the information encoded if each information unit is considered in its own right. To our knowledge, simultaneously encoded distinct information, regardless of its grammatical status, available to the interlocutor has been described only by Dudis (2004)[5] as elaborated in the previous paragraph (see Section 1.2.). Surprisingly, out of all research on amount of simultaneously encoded information, only Risler (2007) explicitly acknowledges the central role of iconicity by accentuating the importance of the diagrammatic link between arrangement of articulators and their semantic representation.

To summarize, diagrammatic iconicity constitutes a peculiar tool in sign languages for packaging semantically related information which is particularly useful when dealing with highly informative events involving multiple referents and their interaction. Instead of encoding each information unit sequentially, signers can construct it as a single representation or superimpose new information on already encoded information. As a result, multiple related information units are explicitly accessible in encoding and decoding. It is then logical to assume that more direct mapping of the event through simultaneity would lead to more efficient encoding in comparison to strictly sequential encoding. However, while previous research provides understanding of how simultaneous information encoding can be achieved and suggests that it might be used for efficiency and informativeness, there is no direct experimental evidence that this peculiar property of sign languages to encode related information simultaneously is actually used for such purpose.

2. Present study

One would expect that if clustering related information lies at the heart of efficient communication, language should adapt to promoting related information being found close together (Hawkins, 2004). Up to now it has been studied how this is achieved in spoken languages (i.e., reduction of dependency distances) but not in signed languages. Exploring how simultaneous encoding of information is used in a systematic way could shed the light on whether this encoding strategy is used in sign languages for efficient communication. While there is some descriptive research available on how simultaneous encoding of information is achieved and how much information can be encoded simultaneously as discussed above, it is not clear when signers would use simultaneity of encoding.

In the present study, we aim to test whether signers use simultaneous information encoding as a strategy to achieve efficient communication, specifically in LIS. We hypothesize that the more information related to the same event that is perceived simultaneously in the world has to be encoded, more articulators encoding different information will be used simultaneously to promote clustering related information closer together. Furthermore, we hypothesize that the more information has to be encoded, the denser simultaneous encoding will be used (i.e., more

information units encoded simultaneously) in order to achieve more direct representation of the event. We test our hypotheses by means of elicited task paradigm in which we assess how amount of information that has to be encoded influences the amount and density of simultaneity of encoding in deaf signers of Italian Sign Language (LIS) in strictly informative context (as opposed to narrative context which requires not only referential but also evaluative function). Considering that previous research on simultaneity has mainly described simultaneous encoding of events involving actions of animate referents, we assess how signers use simultaneous information encoding when dealing with the same kind of events. However, unlike extracting examples from narrative data as has been done before, in which simultaneous constructions can be driven by many different factors and which are not directly comparable to each other, we construct single image stimuli that vary systematically in regard to how many information units they contain. We then present these stimuli to the participants in a context of a game in which they have to describe the images in order for the other person to choose the correct image. Accordingly, in the present task signers are faced with necessity to be as informative and as clear as possible but without additionally enhancing the description with evaluative properties typical for narratives. As a result, we can directly test how signers manipulate the use of simultaneous information encoding and whether they strive for simultaneity when faced with a task of encoding an event that in its nature is perceived simultaneously in the world.

2.1. Method

The study has been approved by the Ethics Council of the National Research Council of Italy (protocol n. 0012633/2019).

2.1.1. Participants

Twenty-three deaf adults (12 female, M age = 30.5, range 18–57) participated in the study. Sixteen participants were native signers of LIS, all children of deaf parents. Seven participants were children of hearing parents and acquired LIS between ages 4–8 at school. All participants reported using LIS daily as the main language. Given some differences in regard to age of acquisition of the participants, we account for it in the analyses. Participants were recruited via mailing list available to The Institute of Cognitive Sciences and Technologies and via recruitment video created in LIS posted on various social media sites. All participants signed consent forms agreeing to be video-recorded and consenting that their data could be used for academic and scientific purposes. For their participation, participants received 5 EUR.

2.1.2. Design

In our design, we systematically increased the information density[6] of the messages that have to be encoded. In the first level only two referents (i.e., two information units) had to be encoded. Participants could encode them by positioning referents in space or also by simply naming them. In second level, participants had to encode both referents and one static action; in third level - two referents, one static action and one dynamic action of the same referent (action1); in fourth level - two referents, one static action of one referent and one dynamic action of the other referent (action2). Finally, in fifth level the participants had to encode two referents, one static action and two dynamic actions of both referents.

---

[4] Process signs are defined as "signs that express processes (actions or events)" (Risler, 2007, p.73)

[5] In his article, Dudis (2004) follows the theory of "conceptual blending" by Fauconnier and Turner (1996) and accordingly discusses amount of visible distinct elements projected into real-space blends in signed production (p.225). Considering that we cannot elaborate on this theory in the article, we rephrase the phenomenon described by Dudis as "simultaneously encoded information".

[6] *Information density* in this study is quantified as total number of information units per experimental item. The term *information* in this study is used to refer to distinct semantic meaning units. Accordingly, we do not draw any parallels with information density of Uniform Information Density framework (Jaeger, 2006; Jaeger & Levy, 2007), where information is defined in "information-theoretic sense—the negative log-probability of an event [..]" (Jaeger & Levy, 2007, p.849) and information density defined as "the amount of information per unit comprising the utterance" (p.849).
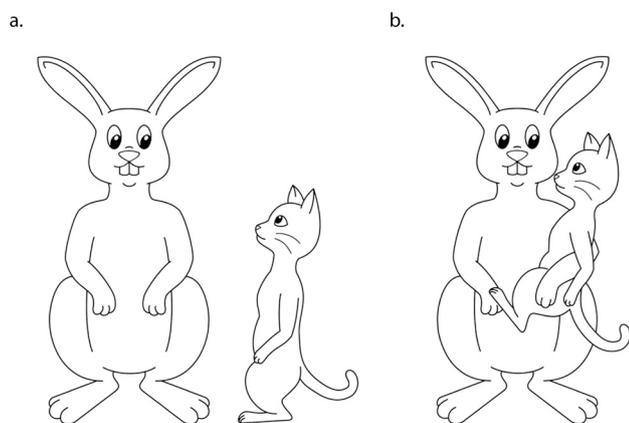
Fig. 3. PNG stimuli of density level 1 (a) and level 2 (b).

Density Level 1 (two referents = two information units in total) (Fig. 3a):

The least dense scenario in our design required referring to two animate referents. For example, a bunny and a cat. This information constituted the first density level in the design where two information units were required to be encoded.

Density Level 2 (two referents + 1 static action = three information units in total) (Fig. 3b):

Then, we increased density of information by one unit by attributing a static action of holding to one referent in relation to the other referent that is being held. For example, the bunny holding the cat. Accordingly, in Density level 2, three information units were required to be encoded.

Density Level 3 (two referents + 1 static action + 1 dynamic action = four information units in total) (Fig. 4 a):

Next, we increased the information density by one more unit and added a dynamic action to the referent doing the static action. For example, the bunny holding and petting the cat. This results in 4 information units that had to be encoded.

Density Level 4 (two referents + 1 static action + 1 dynamic action of other referent = four information units in total) (Fig. 4b):

Then, we shifted the agent of the dynamic action and e.g., instead of the bunny holding and petting the cat, we had the bunny holding the cat and the cat tapping the cheek of the bunny. In this level the bunny became not only the agent but also the patient of the action. Note that here the amount of density did not vary from the previous scenario (we took away one action from one referent and added to another).

However, here characters were both agents and patients simultaneously, thus creating perceptually more complex event and a need to refer to both referents in order to encode both actions in contrast to the previous level were only one referent of the action had to be identified. Accordingly, while we did not manipulate information density between levels 3 and 4, we manipulated only the complexity between these two levels.

Density Level 5 (two referents + 1 static action + 2 dynamic actions = five information units in total) (Fig. 4c):

Finally, we increased the density of the message even further by including both dynamic actions of the referents, resulting in e.g., the bunny holding and petting the cat while the cat is tapping the cheek of the bunny. Here 5 information units had to be encoded.

In encoding of all levels, signers could use different strategies of information encoding - sequential, simultaneous and mixed. Note, that it is impossible to encode the message without first introducing the referents via lexical signs, thus resulting in presence of linearity on all levels. However, the encoding of the event itself is more flexible and could potentially lead to both sequential and simultaneous encoding.

Note that in all the stimuli the eye-gaze of the referents was always kept constant - both referents were looking at each other throughout.

2.1.3. Material

The material for the experiment consisted of 6 sets of 5 stimuli (PNG images for levels 1 and 2 and GIFs for levels 3, 4, and 5) representing each information density level in each set (see Figs. 3 and 4). We decided to use GIFs for the stimuli that involved dynamic actions based on the pilot studies which revealed deaf signers having difficulty interpreting the movement of the dynamic actions from still drawings.

Total of 30 unique combinations of referents and actions were prepared for the experiment (see Appendix A). In each experimental trial, two animate referents were represented. In order to reduce bias for signers to personify with animate referent that is a person, both referents were animals with exception of one pair in which both referents were humans (a woman and a boy). In picture sets, animals alternated between referent 1 and referent 2. Referent 1 was always the bigger referent represented on the left side of the image. Referent 2 was the smaller animal represented on the right side of the image.

2.1.4. Procedure

A participant was greeted by an experimenter (a deaf researcher) and informed that the participant is about to play a *director-matcher* type game with another deaf person. The experimenter also noted, that all instructions would be given via video-recording once the experiment started. The participant was standing in front of another player seated in
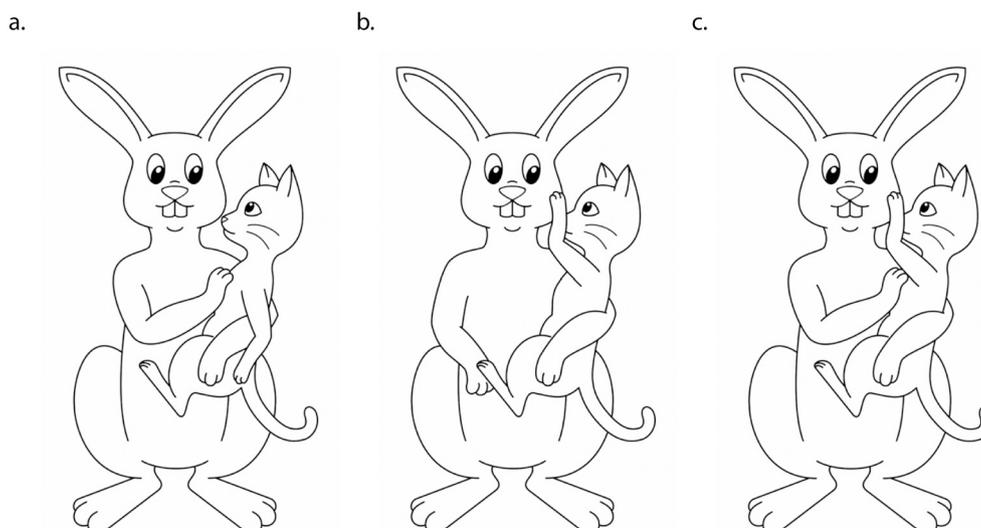

Fig. 4. GIF stimuli of density level 3 (a), level 4 (b), and level 5 (c). In these GIFs dynamic action of the referent 1 and dynamic action of the referent 2 are animated.
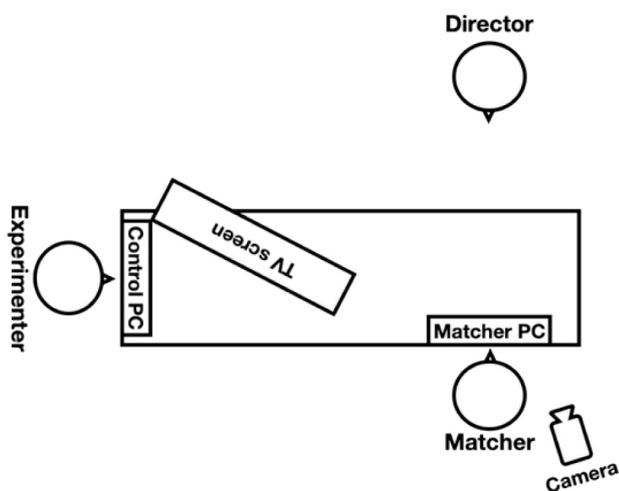
Fig. 5. Experimental setup. Camera located on the tripod (approx. height - 1,50 m).

a chair in front of a table with a laptop (see Fig. 5). This person was a confederate and not a naïve player. It was necessary to use a confederate for the task in order to insure that all participants would receive comparable feedback. Previous research has shown that participants tend to adjust their communicative strategy based on the interlocutor's feedback (Holler & Wilkin, 2011). Confederate was instructed to always provide positive feedback (i.e., a head nod, *OK* sign, *yes* sign, *got it* sign) and no signals of doubt once participant has finished description of the stimuli.

To the left side of the table there was a 40-inch screen in which all instructions and stimuli were presented and which was not visible for the matcher. The experimenter was seated at the left side of the table and controlled the presentation of the video and stimuli by means of the laptop connected to the TV screen.

When participant was standing in the right place the experimenter started instruction video. The participant received an instruction (provided on a video with a recording in LIS) that they were about to play a *director-matcher* game and that they were assigned the role of *the director*. They were instructed that *the matcher* had multiple images presented on their laptop. The task of *the director* was to describe images so that *the matcher* could choose the right one. Participants were asked to look carefully to the screen, memorize the image/GIF and then sign it to the addressee. Pilot study revealed that sometimes it was difficult for the participants to remember the images, thus we decided to leave images on the screen in case there was a need to double-check for the participants. However, they were explicitly asked not to look back at the screen while signing but face the addressee in order to make comprehension easier for the other player.

First, participants were informed that before the start of the actual game they would first see all the characters that have been picked for the game (i.e., the referents) and that they should name these characters to *the matcher* to make the guessing of the characters in actual game easier. Thus, participants could describe each referent separately where they could also be able provide details about their looks and physical aspects and introducing them to the addressee. We chose to present all referents separately first with a consideration that this task would make it less likely that participants would concentrate on aspectual details of referents during the experiment and will instead concentrate on the focus of the study – the referents and their actions. As a result, we would have cleaner data. Each character was presented via PowerPoint presentation one by one. When *the director* had named a character and *the matcher* had nodded to have understood and have picked the image on his laptop the experimenter proceeded to the next image. This procedure also provided a warm-up session and grasp of the game.

Once all characters have been named, experimenter announced that all the characters of the game have now been revealed and the game itself can start. If the participant had no further clarification questions, the experiment started. The process was the same as with the naming images with single characters. If the participant had omitted an action (e.g., in density level 4, the participant forgot to mention that referent 1 is holding referent 2 but only encoded the dynamic action of referent 2), the experimenter asked them to watch carefully the stimuli again and repeat the production. All stimuli were presented in a semi-randomized order. We randomized stimuli in such way that the same referent pair did not appear one after another, also the same density level did not appear one after another. Accordingly, it was necessary to encode all information units depicted in the images rather than contrasting only specific features. No contrasting strategy appeared in our data and each image was always described independently from other images. Productions of the participants were video-recorded and used for coding.

2.1.5. Coding

The video-recorded data was coded in multimodal data annotation software ELAN developed by Max Planck Institute for Psycholinguistics (Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006).

For the present study we developed the coding scheme in order to assess the simultaneous productions of the participants. For each stimulus (annotated with its Density level and Character pair) we coded:

Length of the encoding - Total number of "movement segments" per each production.

Kinematic simultaneity - Simultaneous versus non-simultaneous use of manual and/or non-manual articulators in each movement segment.

Density of simultaneously encoded information - Total number of encoded information units in each movement segment:

1 unit (e.g., Referent 1).

2 units (e.g., Referent 1 + Referent 2).

3 units (e.g., Referent 1 + Referent 2 + Dynamic action of Ref.1).

4 units (e.g., Referent 1 + Referent 2 + Dynamic action of Ref.1 + Static action of Ref.1).

5 units (e.g., Referent 1 + Referent 2 + Dynamic action of Ref.1 + Static action of Ref.1 + Dynamic action of Ref. 2).

2.1.5.1. Length of the encoding and movement segments. To determine sequential organization of the production we segmented data into movement segments based on the start and end of a movement of the hand/s: i.e., segmentation of *a stroke* in terms of Kendon (2004). However, in our coding (as different from Kendon) a movement segment could include not only the stroke produced by the hand but also the hold of the previous sign if present during the new stroke (see Fig. 6, MS 3–5). Thus, movement segment is based on the changes in (at least one) hand movements. Also, if two hands produced independent signs simultaneously (e.g., *holding* with left hand and pointing to *the bird* with right hand, see Fig. 6, MS 2) it was annotated as a single movement segment. Additionally, marked non-manual articulators (change in torso position, change in head position, face expression, eye gaze direction) in each movement segment were annotated.

For example, in Fig. 6, a signer encoded a stimulus from information density level 4 with characters being a dog and a bird. The signer first introduced a dog by means of lexical sign (LS) *dog* in MS1. Then, in MS2 the signer pointed to himself (i.e., signaling assuming the role of the dog) and simultaneously encoded action of holding through constructed action (CA). He then maintained action of holding throughout the production while introducing the bird by first pointing to it (MS 3) and then producing lexical sign for *bird* (MS 4). Note that in MS 3 head and eye-gaze direction was referential as it often occurs together with index signs (Engberg-Pedersen, 2003). Accordingly, change in these non-manual parameters did not indicate marking of the referent but instead reinforced the index sign to the other referent being held. In the last movement segment, while maintaining the holding action, the signer

| MS1	MS 2	MS 3	MS 4	MS 5
RH: LS (dog)	INDEX: self (dog)	INDEX (bird)	LS (bird)	LS (peck)
LH:	CA (holding)	CA (holding)	CA (holding)	CA (holding)
Non-manuals:				(head-dog, face-bird)
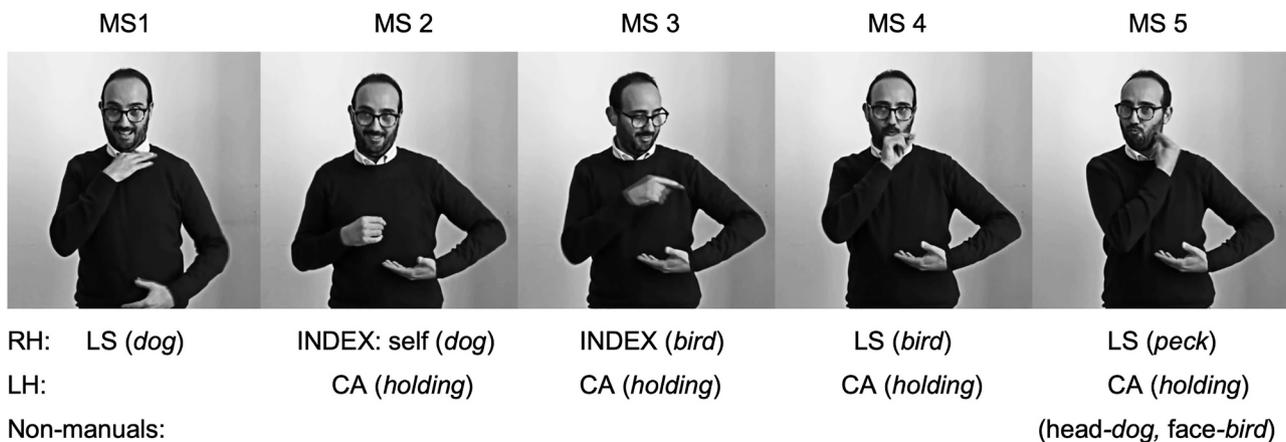
Fig. 6. Example of the segmentation of the movement segments of a single stimulus (Level 4, referents: dog and bird): 5 movement segments in total.

encoded the bird *pecking* through lexical sign (MS 5). In MS 5, the dog was marked via the head of the signer as receiver of the action of the bird. The bird, instead, was marked through the mapping of the bird's mouth onto the mouth of the signer.

It was crucial for us to start with movement segments as the base level, given that we then could proceed to unpack each segment in regard to whether it contained kinematic simultaneity and how many information units were encoded simultaneously in a single segment of time. As such movement segments constituted the sequential nature of encoding. Once we have segmented data on a linear scale (sequentiality of movement segments) we then proceeded to assess whether these segments contained simultaneity (kinematic simultaneity and density of simultaneously encoded information).

We excluded all movement segments that were clear disfluencies or mistakes after which signers corrected themselves. Also, given that we focused on how signers encoded 5 information units that we manipulated in different levels, we excluded additional movement segments that added extra information that was not the focus of our study (e.g., size or shape of the referents, movement segments encoding only eye-gaze direction of referents).

2.1.5.2. Kinematic simultaneity. We then coded whether articulators (manual and non-manual: left-hand/right-hand/torso/head/eye-gaze/face expression) were used simultaneously to encode different information units (Referent1, Referent2, static action, dynamic action1, dynamic action2) in each movement segment. If more than one articulator was used to encode different information units, the movement segment was coded as "simultaneous". Accordingly, this coding showed how many movement segments in each production contained kinematic simultaneity.

2.1.5.3. Density of simultaneity. In a separate tier, we counted how many information units were encoded simultaneously within each movement segment. In the design, we constructed stimuli to focus on following information units: referent 1, referent 2, static action, dynamic action of referent 1, dynamic action of referent 2. This sums up to 5 information units that had to be encoded in our design, amount of which was dependent on the density level of the stimuli. In each movement segment, we counted how many information units of interest were simultaneously and explicitly available to the interlocutor. Thus, implicit referents (e.g., referent 2 implicitly available by the form of holding action, see Fig. 7) were not counted.

When encoding action by hand, referent had to be marked by at least one non-manual marker (eye-gaze, face expression, head, torso) in order to be counted as encoded (e.g., see Figs. 7, 8, 9). In cases where two actions of both referents were encoded simultaneously, availability of the second referent was only coded if it was explicitly marked by eye
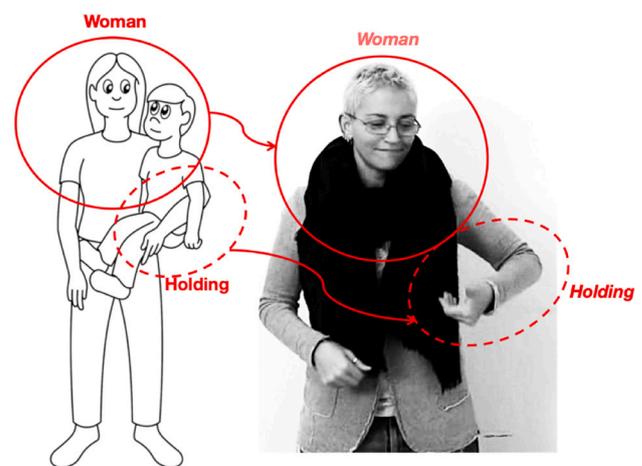

Fig. 7. Movement segment of simultaneous encoding of *2* information units - Referent 1 (encoded through head direction, face expression and eye-gaze) and the static action (the left hand of the signer).
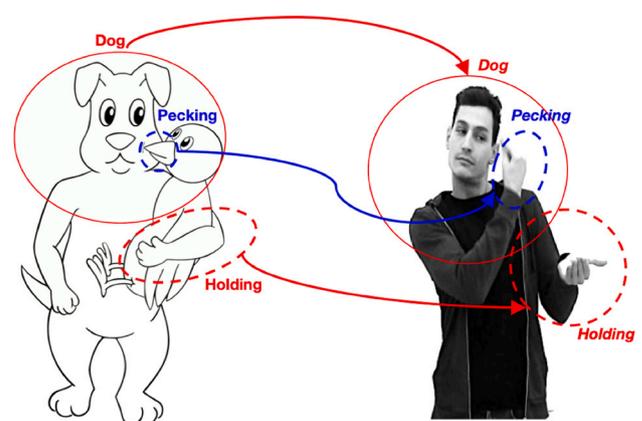

Fig. 8. Movement segment of simultaneous encoding of 3 information units - Referent 1 (encoded through torso, head, eye gaze and face expression of the signer), static action (left hand of the signer) and dynamic action of Ref.2 (lexical sign of *pecking* signed by the right hand).
Stimulus from the information density level 4.

gaze, head, or face expression (Fig. 9). If not, we coded that only the action of the referent2 was encoded (Fig. 8).

Stimulus from the information density level 2.
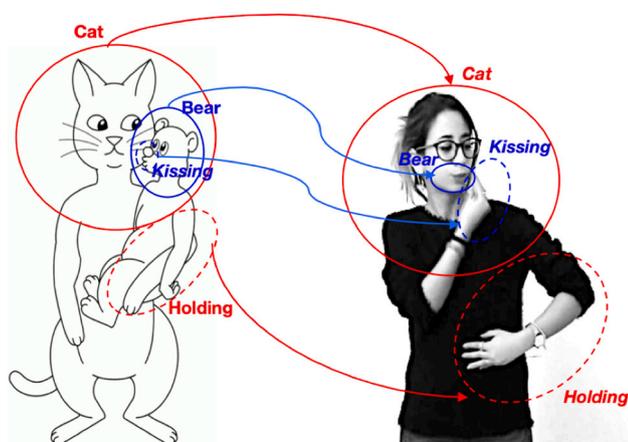
Fig. 9. Movement segment of simultaneous encoding of 4 information units - Referent 1 (encoded through torso, head and eye gaze of the signer) and static action (left hand of the signer), Referent 2 (encoded through mouth) and dynamic action of Ref.2 (lexical sign of *kissing* signed by the right hand). Stimulus from the information density level 4.

2.1.6. Reliability

All data was initially coded by the first author of the study. All coded data was double-checked by a deaf researcher, a native signer of LIS. Another native signer of LIS independently coded 20% of the data.

Agreement between coders was almost perfect – 98.53% for gross-level segmentation of movement segments. Agreement assessment was done based on Kita et al. (1997, p.10): The gross-level segmentation is recognition of a stretch of movement with a certain directionality [..] as a phase, regardless of the exact location of the boundaries and the identification of the phase type [..]". Out of 885 annotated movement segments coders agreed on 872 movement segments. We then derived reliability statistic by assessing the total number of movement segments per stimuli. It was very strong as revealed by *Cohens* $\kappa = 0.94$.

Reliability of simultaneous use of multiple articulators (manual and non-manual) in each movement segment (*Cohen's* $\kappa = 0.94$) as well as reliability regarding number of simultaneously encoded information units in each movement segment (*Cohens* $\kappa = 0.95$) was very strong.

2.1.7. Analyses

We analyzed data in R by using generalized mixed models (package *lme4*, Bates, Maechler, Bolker, & Walker, 2015). The significance of the mixed model were derived from model comparisons starting with the baseline model that includes random effects and significant confounding factors. Significance levels were derived by using Satterthwaite's method provided by R package *LmerTest* (Kuznetsova, Brockhoff, & Christensen, 2015). We use this method as it is proven to be the most conservative in regard to Type I errors, i.e., false positives (Luke, 2017).

We used the method of mixed effects models to test the effect of the *Information Density Level* on the *length of the production, kinematic simultaneity, density of simultaneity* used. Mixed effects models make possible examining not only the fixed effects, but also include random effects of the individual trials and participants. Moreover, mixed effects models allow modeling not only random intercepts but also random slopes and thus can account for even more fine-grained individual variation that might have influence on the outcome of the analyses.

In our study, the following random effects were considered for the model: *trial (stimuli sample), participant and character pair*. By including random intercept (i.e., random effect) for the stimuli sample and character pair we account for possible variability that some specific stimuli or specific character pair might be generally more powerful in eliciting simultaneity from the participants than others. Also, some participants might be in general more prone to use simultaneity than other participants, thus random intercept for participant was used. It is also possible

that the effect of the density level is stronger for some participants and not for the others. Thus, in order to account for this aspect, we also considered random slope of *density level by participant*. Accordingly, the individual differences of participants in regard to how sensitive they are to the predictor variable could be controlled for. Furthermore, we run series of models to account for possible confounding fixed factors, e.g., *gender, age, age of LIS acquisition, handedness*. Final baseline model was determined based on the best fit as revealed by ANOVA tests or alternatively on the maximal random effects structure that converged in the model (Barr, Levy, Scheepers, & Tily, 2013).

All stimuli, data and analyses scripts are available online (https://doi.org/10.17605/OSF.IO/MWG4V).

2.2. Results

In the present experiment we tested whether participants varied amount of kinematic simultaneity and density of simultaneously encoded information based on the information density level they had to encode. The results section is organized as follows. First, we assess what are the differences in regard to the length of the production in each information density level. Next, we test whether kinematic simultaneity (2 or more articulators used simultaneously to encode different information) increases with the increase of the information density level. We hypothesized that participants will increase kinematic simultaneity to encode informatively denser messages. Next, we explore in which levels the most simultaneously encoded information (i.e., number of simultaneously encoded information units in a single movement segment) is used. We hypothesized that as the information that has to be encoded gets denser so does the density of simultaneously encoded information. The results are based on 23 participants describing 30 items varying in their informative density. We excluded 12 out of total of 690 trials in which signers produced an incomplete description (e.g., omitted one or more information units). In order to be more conservative in regard to natural production of the participants, we did not consider corrected descriptions when prompted by the experimenter. Accordingly, the results are based on 678 trials in total.

2.2.1. Length of the production

In Fig. 10 we present raw means of the total number of movement segments (MS) per stimuli used in each information density level. The participants used on average 3.72 MS (SD = 1.04) to encode 2 information units in Level 1; 4.33 MS (SD = 0.73) on average to encode 3 information units in Level 2; 5.60 MS (SD = 1.18) on average to encode 4 information units in Level 3; 5.93 MS (SD = 1.30) on average to encode 4 information units in Level 4; and 7.68 MS (SD = 1.64) on average to encode 5 information units in Level 5. The statistical analysis was based on 678 data points (experimental stimuli/trials).

A generalized mixed effects model was fit to assess the fixed effect of *Information Density Level* (coded as categorical variable with 5 levels) on *the length of the production* quantified as a count of total number of movement segments per stimuli (family = *poisson*). The effect of *Information Density Level* was compared to the baseline model which included random effect of *participant* and random effect of *trial*. We run series of models to test possible confounding factors - *gender, age, age of LIS acquisition, handedness and character pair*. None of the factors was significant. In the remainder of the study we do not mention these factors unless they prove to be significant. There was a significant main effect of *Information Density Level* ($\chi2(4) = 83.47$, $p < .001$) (see Table 1).

The primary model was releveled in order to attain hierarchical contrasts between the levels. Pairwise comparisons revealed a significant gradual increase in the length of the production as the information that had to be encoded increased (Level 1 vs. Level 2: $\beta = 0.15$, SE = 0.06, CI[0.03, 0.27], $z = 2.47$, $p = .01$; Level 2 vs. Level 3: $\beta = 0.26$, SE = 0.05, CI [0.15, 0.37], $z = 4.72$, $p < .001$; Level 4 vs. Level 5: $\beta = 0.27$, SE = 0.05, CI[0.17, 0.36], $z = 5.62$, $p < .001$) except for levels 3 and 4 for which the difference was only borderline significant ($\beta = 0.05$, SE = 0.05, CI[−0.04, 0.15], $z = 1.07$, $p = .28$).
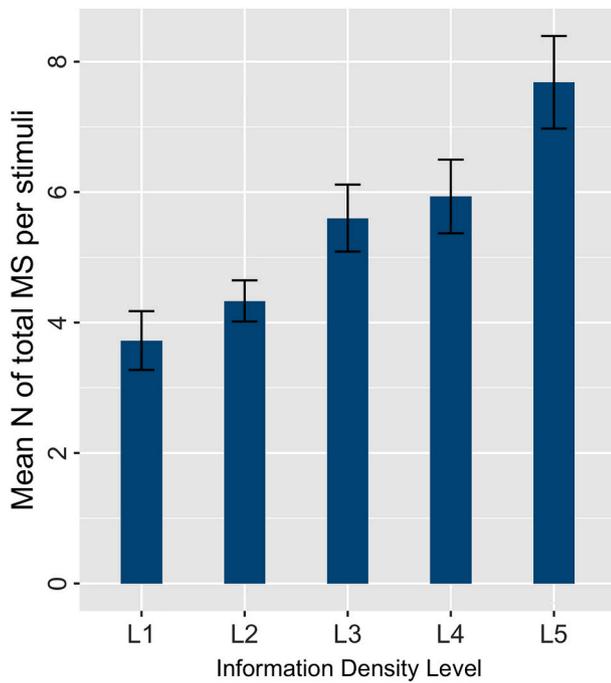
Fig. 10. Mean of total number of movement segments used per each stimuli. Error bars indicate 95% CI of observations grouped within participants.

2.2.2. Simultaneity

2.2.2.1. Kinematic simultaneity.
Considering the differences in the length of the production, analysis was based on proportions of movement segments (MS) with kinematically simultaneous articulators (i.e., two or more articulators used in a single movement segment) out of total number of movement segments per stimuli (see Fig. 11). The movement segments expressed with kinematically simultaneous articulators were used scarcely (M = 0.05, SD = 0.07) in the least dense information level (Level 1). In Level 2 almost half of the MS (M = 0.46, SD = 0.16), contained kinematically simultaneous articulators. Kinematic simultaneity increased further as the information density that had to be encoded increased: Level 3 (M = 0.57, SD = 0.12), Level 4 (M = 0.60, SD = 0.10),

Table 1
Best fit model in a logit scale (model fit by maximum likelihood, Laplace approximation) regarding use of total number of MS per experimental stimuli. Contrasts reflect pairwise comparison between Level 1 and all other levels.

Random effects	Variance	SD
Trial	0.00[a]	0.00
Participant	0.02	0.15
Number of obs: 678	Groups: Trial = 30,	Participant = 23

Fixed effects	95% CI		β	SE	z value	p value
	Lower b.	Upper b.				
(Intercept)	1.20	1.41	1.30	0.05	23.94	<.001
Level 2	0.03	0.27	0.15	0.06	2.47	.01
Level 3	0.30	0.52	0.41	0.06	7.18	<.001
Level 4	0.35	0.57	0.46	0.06	8.18	<.001
Level 5	0.62	0.83	0.73	0.05	13.47	<.001

[a] Zero variance of the random effect of trial is driven by inclusion of the fixed effect of Information Density Level, which accounts for all variance detected in random effect of trial in baseline model. Given that inclusion of trial is based on the initial design of the study and the results do not change if this random effect is left out, we keep it in the primary model. Controls of random effect of trial can be found in supporting material. This consideration applies to all consecutive analyses.
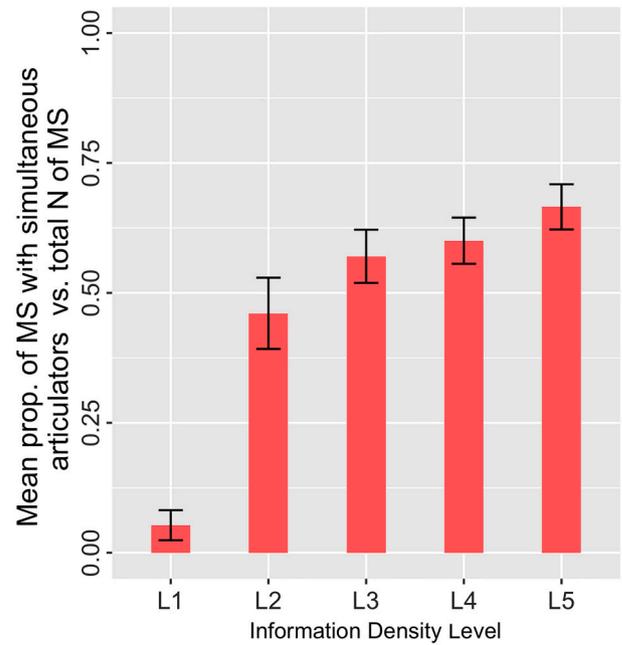

Fig. 11. Raw mean proportions of kinematically simultaneous movement segments out of total number of movement segments per stimuli. Error bars indicate 95% CI of observations grouped within participants.

and Level 5 (M = 0.67, SD = 0.10). The statistical analysis was based on 678 data points (experimental stimuli/trials).

A generalized mixed effects model was fit to assess the fixed effect of *Information Density Level* on the amount of *kinematic simultaneity* used quantified as total number of MS containing simultaneity versus total number of MS used to encode each trial (family = *binomial*). The effect of *Information Density Level* was compared to the baseline model, which included random effect of *participant* and random effect of *trial*. The main effect was highly significant ($\chi2(4) = 110.16$, $p < .001$) (see Table 2).

The primary model was releveled in order to attain hierarchical contrasts between the levels. Pairwise comparisons revealed that there was a significant gradual increase in use of MS with kinematically simultaneous articulators (Level 1 vs. Level 2: $\beta = 2.61$, SE = 0.20, CI[2.22, 3.00], $z = 13.19$, $p < .001$; Level 2 vs Level 3: $\beta = 0.47$, SE = 0.11, CI[0.25, 0.69], $z = 4.15$, $p < .001$; Level 4 vs. Level 5: $\beta = 0.33$, SE = 0.10, CI[0.13, 0.52], $z = 3.26$, $p = .001$) except for Levels 3 and 4 which were comparable ($\beta = 0.08$, SE = 0.1, CI[−0.12, 0.29], $z = 0.77$, $p = .44$).

Table 2
Best fit model in a logit scale (model fit by maximum likelihood, Laplace Approximation) regarding the proportion of kinematically simultaneous MS. Contrasts reflect pairwise comparison between Level 1 and all other levels.

Random effects	Variance	SD
Trial	0.00	0.00
Participant	0.15	0.39
Number of obs: 678	Groups: Trial = 30,	Participant = 23

| Fixed effects	95% CI		β	SE	z value	p value
	Lower b.	Upper b.				
(Intercept)	−3.12	−2.35	−2.74	0.20	−13.91	<.001
Level 2	2.22	3.00	2.61	0.20	13.19	<.001
Level 3	2.70	3.46	3.08	0.19	15.87	<.001
Level 4	2.78	3.54	3.16	0.19	16.30	<.001
Level 5	3.11	3.86	3.49	0.19	18.17	<.001
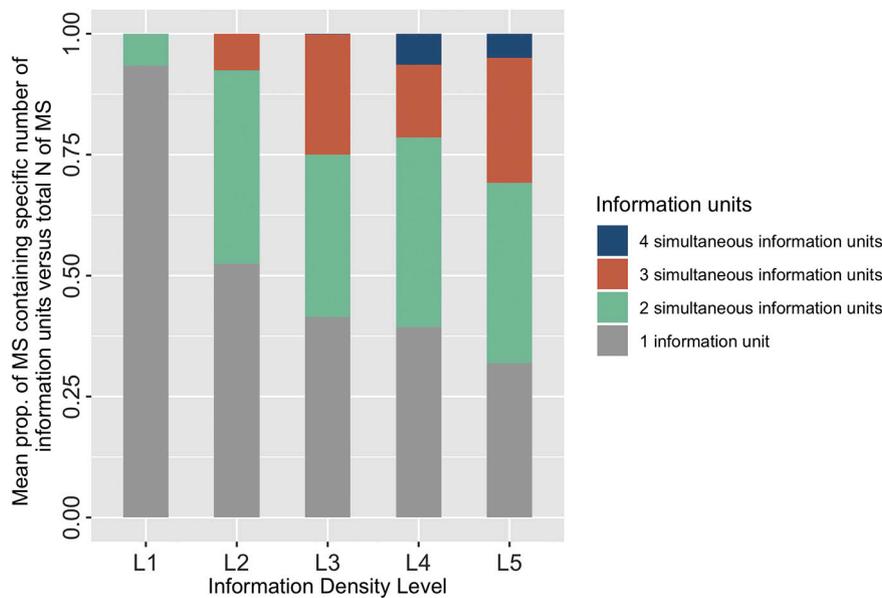
Fig. 12. Raw mean proportions of MS with 1, 2, 3, and 4 simultaneous information units out of total number of movement segments.

2.2.2.2. Density of simultaneously encoded information. Overall, there were 3697 movement segments used out of which 1748 MS with one information unit, 1225 MS with two information units, 622 MS with three information units, and 102 MS with four information units. Accordingly, the data revealed that the maximum of four information units were encoded simultaneously (see Fig. 12). However, they were found only in Information Density Levels 4 and 5 (except one instance of four information units in Level 3. MS with two and three simultaneous information units were found in all levels. Note that in Level 1 all kinematically simultaneous movement segments contained exclusively 2 information units. The statistical analysis was based on 3697 data points (movement segments).

In order to test whether density of simultaneously encoded information increased based on the increase of Information Density Level of the trial, we run a generalized mixed effects model where the fixed effect was *Information Density Level* and the outcome variable was the *density of simultaneity* quantified as count of total number of information units encoded in a single MS (family = *poisson*). The effect of *Information Density Level* was compared to the baseline model, which included random effect of *participant* and random effect of *trial*. The main effect was highly significant ($\chi2(4) = 88.74$, $p < .001$) (see Table 3).

The primary model was releveled in order to attain hierarchical contrasts between the levels. Pairwise comparisons revealed that as

Information Density Level increase so did the Movement segments used to encode these levels become denser in regard to information they contained. (Level 1 vs. Level 2: $\beta = 0.38$, SE = 0.05, CI[0.27, 0.48], $z = 6.92$, $p < .001$; Level 2 vs Level 3: $\beta = 0.17$, SE = 0.04, CI[0.09, 0.25], $z = 4.00$, $p < .001$; Level 4 vs. Level 5: $\beta = 0.08$, SE = 0.03, CI[0.01, 0.14], $z = 2.29$, $p = .02$) except for Levels 3 and 4 which were comparable ($\beta = 0.02$, SE = 0.04, CI[−0.05, 0.1], $z = 0.68$, $p = .50$). It appears that the lack of difference in increase of density of simultaneously encoded information is driven by signers using more MS with 2 information units in Level 4 than in Level 3. As a result, the incremental use of MS with 4 information units is balanced out at expense of higher use of MS with 2 information units.

3. Discussion

In the present study we hypothesized that as the information to be communicated increases, so will use of kinematic simultaneity and density of simultaneously encoded information. Our data confirmed that not only the length of production increased but also use of kinematic simultaneity. Furthermore, as the information to be communicated became denser, so did movement segments (MS). These findings indicate that at least in LIS signers use simultaneous information encoding to achieve efficient communication.

Given that not only simultaneity but also the length of production increased following informative demands indicates that simultaneity was not simply used to reduce the number of movement segments (which might seem a straight forward prediction), but it was used for simultaneous encoding of multiple related information units. Increase in both length and simultaneity might lie in the contrasting needs to be both informative and brief (Grice, 1975; Perniss, 2007a, 2007b). Presenting an entire utterance in a single simultaneous construction right away might lead to ambiguity, while stacking subsequent one or multiple units on existent information may accommodate for informative and temporal needs. Alternatively, presence of the increase in linearity might also be linked to motoric constraints (e.g., producing two actions simultaneously requires more planning, (Oliveira and Ivry, 2008) or absence of time constraints (signers were not under pressure to sign as fast as possible). However, systematic comparisons with spoken languages are needed, to fully understand the relation between simultaneity and length of production in signed languages. So far comparisons of information encoding in spoken and signed languages have shown

Table 3
Best fit model in a logit scale (model fit by maximum likelihood, Laplace Approximation) regarding use of MS with increasing information density (N of information units MS contains). Contrasts reflect pairwise comparison between Level 1 and all other levels.

Random effects	Variance	SD
Trial	0.00	0.00
Participant	0.001	0.03
Number of obs:3697	Groups: Trial = 30,	Participant = 23

Fixed effects	95% CI		β	SE	*z* value	*p* value
	Lower b.	Upper b.				
(Intercept)	−0.02	0.15	0.06	0.04	1.43	.15
Level 2	0.27	0.48	0.38	0.05	6.92	<.001
Level 3	0.45	0.64	0.55	0.05	10.84	<.001
Level 4	0.47	0.67	0.57	0.05	11.43	<.001
Level 5	0.55	0.74	0.65	0.05	13.49	<.001

that in cases in which the amount of information is kept constant fewer consecutive linguistic units are needed in signed languages compared to spoken languages (Bellugi & Fischer, 1972). However, further studies are needed to evaluate whether this would also be true in cases in which the density of information to be conveyed grows, such as in the present study. We expect that there would be an increase in linguistic units in signed and spoken languages, but that the former would be less than the latter. A systematic comparison assessing this prediction should be an endeavour for further research.

In regard to density of simultaneity, MS containing two information units were used in all levels, use of MS with three information units were used only in levels 3, 4 and 5 considered in our study; MS with four information units were used only in levels 4 and 5 (with exception of one instance in level 3). Interestingly we never found five information units encoded simultaneously, although, at least for half of the stimuli in levels 4 and 5, this would be possible given that the action performed by the referent2 was done with the face (i.e., beak, tongue, lips). One could have expected that if direct mapping is the aim of the encoding, we would find use of non-manual articulators to encode the third action as well. However, data suggests that this is not the case and, at least in LIS, actions are encoded also using the hands. An alternative explanation, however, might stem from cognitive constraints. Research shows that there is a limit on visual memory capacity which appears to be exactly four elements/objects (Cowan, 2001, 2010; Irwin & Andrews, 1996; Luck & Vogel, 1997; Sperling, 1960; Vogel, Woodman, & Luck, 2001). This is also in line with findings from Napoli and Sutton-Spence (2010) who find a maximum of 4 propositions in ASL. However, research on visual working memory capacity disregards the fact that diagrammatic iconicity may bind multiple elements into single representations. In this respect, it is not clear whether each information unit encoded simultaneously in sign language should be treated as an independent element, and thus the limit of maximum 4 elements would apply, or instead as a feature of a single more complex element.

Additionally, we found that in levels 3 and 4 in which we maintained the same number of information units but manipulated the complexity of the event (i.e., one agent versus two agents), the quantitative increase in use and density of simultaneity was not significant. Signers encoded these two events using the same proportion of MS containing kinematic simultaneity by manipulating the density of the simultaneously encoded information in qualitative manner (i.e., using more MS containing 2 and 4 simultaneous information units in Level 4 in comparison to Level 3). Although assessment of the specific information encoded in MS was not the goal of our study, we speculate that it is conceptually easier to encode a single referent and his actions simultaneously as opposed to simultaneous encoding two referents and their actions, as the latter requires splitting the body in two conceptually distinct entities through body partitioning. Results also indicate that it is possible that for some signers the ease of encoding might influence their striving for maximum simultaneity. Namely, instead of always encoding four information units in the same MS, signers could instead encode two separate MS each encoding one referent and his/her action. Accordingly, our choice to add Level 4 that increases complexity while maintaining the same number of information units as in Level 3 seems justified as it provides some preliminary insights into how use of simultaneity is affected by the complexity of the event while informative density is held constant. In line with results discussed in the previous paragraph, in the future it would be necessary to explore whether simultaneous encoding is constrained by cognitive demands and/or by articulatory effort.

3.1. Simultaneity over linearity for efficient communication in sign languages

Our results show that as the information that had to be encoded increased, signers increased amount and density of simultaneously encoded information, which was only possible due to the bounding property of diagrammatic iconicity. In other words, diagrammatic iconicity allowed signers to intertwine additional information to the diagram resulting in a single representation. If anything, without diagrammatic iconicity, simultaneous encoding of multiple information units, with no iconic relation connecting them, would hinder efficient communication. We argue that signers in our study strived for a more direct mapping of the meaning and linguistic encoding in order to boost more efficient representation formation. This interpretation appears to be also in line with Chistiansen & Chater (2016) who argue that given the constraints of perceptuo-motor processing and memory "language system engages in eager processing" (p.5) in order to create higher-level linguistic representation as fast as possible. Namely, low-level information is passed onto higher representation levels, e.g., sounds ➔ words ➔ discourse, to form a single chunk of representation which can be retained more efficiently. Note that for spoken languages the low-level chunks (i.e., sounds) are passed onto higher levels in a linear manner. For signed languages, on the other hand, chunking can occur both linearly, i.e., chunking one sign after another, and simultaneously due to diagrammatic iconicity, that relates multiple meaning elements that are produced with different articulators. As a result, use of diagrammatic iconicity would boost the chunking process, considering that higher-level representation could be constructed in a direct relation to other information in a simultaneous manner as opposed sequentially. Signers use mental imagery generation for language production and comprehension (Emmorey, 1993; Emmorey & Kosslyn, 1996; Emmorey, Kosslyn, & Bellugi, 1993) especially if spatial relations are involved (Emmorey, 1995). Thus, advantage of possibility to exploit diagrammatic iconicity to bound different information into single representation might be at play in encoding and decoding, as it can be used to adhere to the mental imagery of the signer and interlocutor. As a result, use of diagrammatic iconicity for simultaneous encoding of information would boost the chunking process, considering that higher-level representation is available right away as opposed sequentially. Indeed, diagrammatic iconicity has been shown to aid conceptual processing also in non-signers (Louwerse & Jeuniaux, 2010; Zwaan, Stanfield, & Yaxley, 2002; Zwaan & Yaxley, 2003), indicating that it can be used as a tool in more efficient representation forming involving multiple elements. For example, Zwaan and Yaxley (2003) showed that simultaneously presented words in reverse-iconic relation (e.g., word *basement* above the word *attic*) resulted in slower semantic-relatedness judgments than words presented in iconic relation (word *attic* above word *basement*). We speculate that exactly simultaneous encoding in signed languages is an efficient strategy to chunk information in higher level representations. Vinson, Thompson, Skinner, and Vigliocco (2015) argue for iconicity as a vehicle for a faster path between meaning and form in production and comprehension on a single lexical sign level. We predict that this finding would also extend to iconic simultaneous constructions. Testing this hypothesis, however would be an endeavour for future research.

The specific sensory constraints present in language production might put specific pressures on how language is organized leading to the exploitation of simultaneity over linearity in signed languages (Supalla, Siple, & Fischer, 1991). Signers are adept at dealing efficiently with the integration of simultaneous information. Especially considering signs' multilinear nature and use of space for grammatical encoding that has to be processed regularly during language use (Wilson & Emmorey, 1997). Indeed, Capirci, Cattani, Rossini, and Volterra (1998) show how even minimal, but constant, exposure to a sign language can boost visual-spatial cognition and spatial memory in hearing children. Therefore, it is not surprising to suggest that when it comes to proficient sign language users, they may take advantage of simultaneous information encoding in order to minimize dependency distances. As a result, it would be less cognitively demanding for signers to encode and decode information presented

simultaneously as opposed to consecutively. Research suggests that when visual working memory is involved it benefits when simultaneity can be taken advantage of (Allen, Baddeley, & Hitch, 2006; Frick, 1985; Wilson & Emmorey, 1997; Woodman, Vecera, & Luck, 2003). The same is not true when processing of auditory information is involved, which seems to be bound instead to a persistent unidirectional effect, i.e., auditory items are recalled in the same sequence as the one in which they have been presented (McFarland & Kellas, 1974; Penney, 1989). Indeed, serial recall is particularly hard for signers in comparison to speakers as indicated by studies documenting shorter memory span in signers (Bavelier, Newport, Hall, Supalla, & Boutla, 2006; Boutla, Supalla, Newport, & Bavelier, 2004; Geraci, Gozzi, Papagno, & Cecchetto, 2008; although see also Rudner & Rönnberg, 2008, and Wilson & Emmorey, 2006). We argue that the finding that signers increase use of simultaneity as information to be conveyed becomes denser, suggests that simultaneity is used strategically to decrease dependency distances and promote more efficient information chunking and, thus, ease processing. For example, when a signer introduces a referent and subsequently becomes the referent in order to encode the action of holding, the referent does not have to be maintained solely in working memory as it is conceptually present in the signer's body. Furthermore, instead of providing each information element one at the time, additional information can be integrated into the iconic diagram to form a more complete representation. Accordingly, to some extent, simultaneous information encoding could function as cognitive offloading and lighten processing (Risko & Gilbert, 2016) considering that some information can be externalized and maintained as new information is introduced. Indeed, also Napoli and Sutton-Spence (2010) suggest that *connectedness* among simultaneous units is a potential candidate for reducing cognitive load (p.675). Whether simultaneous information encoding can indeed lighten cognitive load in sign languages may be tested in the future on both behavioural and neurological levels.

Although we stressed that spoken languages have been described as strictly linear, use of the whole body or of individual bodily parts would also allow speakers to employ bodily articulators to encode distinct information simultaneously (Kendon, 2014). However, it was not our intention to directly compare simultaneous information encoding in sign languages to spoken languages. Accordingly, discussion on multimodality in spoken languages was not elaborated on. While it is expected that signers are more sophisticated in using their body for information encoding, it would be interesting to see whether also speakers take advantage of the visual modality for efficient information encoding through simultaneity (e.g., use of complementary gesture with speech). Previous research indicates that, for the purpose of task-solving, use of gesture with speech can lighten working memory load (Cartmill, Beilock, & Goldin-Meadow, 2012; Cook, Yip, & Goldin-Meadow, 2012; Goldin-Meadow & Beilock, 2010).

Alternatively, it would be important to assess how hearing non-signers can use their body to transmit such information at different Density levels when no linguistic encoding is possible (i.e., no knowledge of sign language). Comparisons between sign languages and silent gesture use can elucidate how much of the iconic structures and simultaneity we find here are due to general visual affordances of using the body and how much due to skillful use of linguistic resources. It is highly possible that more iconicity and simultaneity need complex linguistic tools. Research on these topics where we ask Italian speakers to express same stimuli using silent gestures is currently under way in our lab. Furthermore, future research on grammatical structure would be needed to understand how use of simultaneity interacts with affordances and the constraints of grammatical structure of a specific language. Namely, it would be interesting to explore which information exactly is clustered together and how in both spoken (speech + gesture) and signed languages. In this vein it would be important to investigate specific linguistic strategies used for simultaneous encoding of information. We described in brief how different linguistic strategies allow simultaneity in sign languages, but a systematic assessment of specific strategy use and their combination would be necessary to elucidate how exactly simultaneity is achieved. The investigation on the choice of specific linguistic strategies is currently underway.

3.2. Limitations of the study

There are some limitations to this study that we consider important to stress. First, we were interested in assessing how many information units can be encoded simultaneously in Italian Sign Language (LIS). In our design, we could test simultaneous encoding of two animate referents and their actions. Accordingly, we cannot generalize our findings to overall simultaneity of information, as we explored only a fraction of possible simultaneity in sign languages. Nevertheless, this is the first systematic study to quantify simultaneous encoding of information on events involving animate referents and assess whether this property is used to achieve better communicative efficiency in LIS. Future research should show whether our findings generalize to other sign languages and other types of events as well.

Second, based on our results, we find that out of a maximum of the five information units that we focused on, four units could be encoded simultaneously. Even though we criticized Napoli and Sutton-Spence (2010) for not accounting for the overall simultaneity that signed languages are able to achieve, we may be seen as falling in a similar problem. In fact, even in our design there was more information available to the interlocutor and simultaneously encoded by the participants than what we quantified in our data, (e.g., eye gaze, action location, and size of the referents), we disregarded this information in the present study. There is yet another information unit that we disregarded, namely, the availability of the implicit referent. Dudis (2004) notes that signers can render non-present referents available to the interlocutor via conceptual integration. For example, in instances of signers encoding referent 1 and a static action, the non-present but yet conceptually available referent is the one that is being held, i.e., implicit referent (see Fig. 7). The same applies to instances where signers encode referent 2 and dynamic action2 in a single movement segment. The action in this case is performed on a non-present but yet conceptually available referent 1. In the future, it would be important to account for all the information that is available to the interlocutor in a single movement segment and how it is increased. However, this would require different stimuli and more fine-grained analyses specifically targeting this issue in order to assure an adequate study design. Another interesting question that arises is whether there are differences in cognitive effort when integrating implicitly vs. explicitly encoded information.

4. Conclusion

An event in the world has to travel a long way to reach its encoding in language. Just as in spoken languages also in signed languages signers face the linearization problem (Levelt, 1981). Namely, an event that occurs simultaneously in the world has to be split in language to be organized on a temporal scale. Accordingly, the goal of minimizing dependency distances in order to render encoding more efficient stands also for signed languages. While striving to cluster related information closer together can be considered a general trait of language, due to temporal processing constraints (Christiansen & Chater, 2016), how this is achieved is, to some extent, dependent on the linguistic modality. We hypothesized that in signed languages (i.e., LIS in this specific case) clustering could be achieved by exploiting diagrammatic iconicity to encode multiple related information simultaneously. Signers do not

limit their clustering to relating information consecutively as in spoken languages (Lu et al., 2016), but they encode information simultaneously reducing dependency distances to the minimum. We attempted to quantify how simultaneous encoding of information is used by signers of LIS when encoding messages with various informative density. We found that kinematic simultaneity was used more and density of simultaneously encoded information increased when it was necessary to provide more information.

The core property of iconicity to depict the world as opposed to arbitrariness that describes it "is one of the central affordances of human language" (Dingemanse, 2018, p.19). It reinforces the link between the form and the meaning and allows more direct information transmission (Perniss & Vigliocco, 2014; Vinson et al., 2015). Signers employ iconicity to represent the information present in events as it is available in the real word – simultaneously – and as such they are more truthful to the facts they are referring to. As a result, conceptual representation can be formed faster. We conclude that iconicity in sign languages should be seen as an advantage, as it allows more efficient, i.e., simultaneous, encoding of dense information.

Appendix A. Stimuli used for the experiment

Note that figures in levels 3, 4 and 5 are GIFs and dynamic actions are animated). All images in their original format are freely available online (https://doi.org/10.17605/OSF.IO/MWG4V).

Set: ***Bear*** & ***Dog***. Dynamic action 2 – *licking*.

Set: ***Bear & Dog.*** Dynamic action 2 – *licking*.

Density level 1

Density level 2

Density level 3

Density level 4

Density level 5

Set: **Dog** & **Bird**. Dynamic action 2 – *pecking*.

Set: **Dog & Bird.** Dynamic action 2 – *pecking*.

Density level 1   **Density level 2**
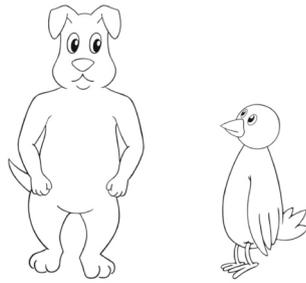
Density level 3   **Density level 4**   **Density level 5**

Set: **Bird** & **Bunny**. Dynamic action 2 – *tapping*.

Set: **Bird & Bunny.** Dynamic action 2 – *tapping*.

Density level 1   **Density level 2**
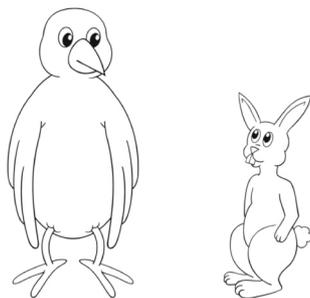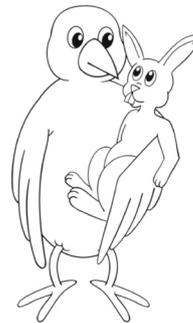
Density level 3   **Density level 4**   **Density level 5**

Set: **Bunny** & **Cat**. Dynamic action 2 – *caressing*.

Set: **Bunny & Cat.** Dynamic action 2 – *caressing*.

Density level 1   **Density level 2**
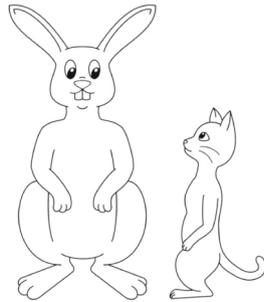
Density level 3   **Density level 4**   **Density level 5**

Set: **Cat** & **Bear**. Dynamic action 2 – *kissing*.

Set: **Cat & Bear.** Dynamic action 2 – *kissing*.

Density level 1   **Density level 2**
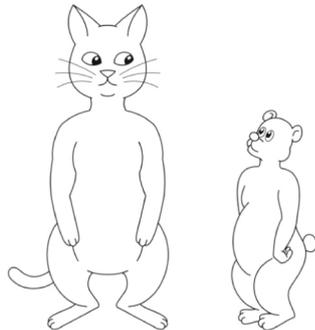
Density level 3   **Density level 4**   **Density level 5**

Set: **Woman** & **Child**. Dynamic action 2 – *pinching*.

Density level 1   Density level 2
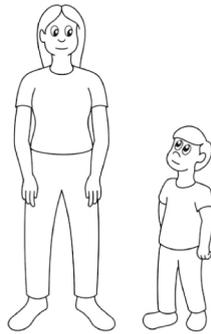
Density level 3   Density level 4   Density level 5

Appendix B. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2020.104246.

References

Allen, R. J., Baddeley, A. D., & Hitch, G. J. (2006). Is the binding of visual features in working memory resource-demanding? *Journal of Experimental Psychology: General, 135*(2), 298.

Aylett, M., & Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech, 47*(1), 31–56.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). *lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1–7. 2014.*

Bavelier, D., Newport, E. L., Hall, M. L., Supalla, T., & Boutla, M. (2006). Persistent difference in short-term memory span between sign and speech: Implications for cross-linguistic comparisons. *Psychological Science, 17*(12), 1090–1092.

Bellugi, U., & Fischer, S. (1972). A comparison of sign language and spoken language. *Cognition, 1*(2–3), 173–200.

Boutla, M., Supalla, T., Newport, L., & Bavelier, D. (2004). Short-term memory span: Insights from sign language. *Nature Neuroscience, 7*, 997–1002.

Boyes Braem, P. (1980). *Significant features of the Handshape in American sign language.* University of California, Berkeley: Unpublished doctoral dissertation.

Bybee, J. L. (1985). Morphology: A study of the relation between meaning and form (Vol. 9). John Benjamins Publishing.

Bybee, J. L. (2013). *Usage-based theory and exemplar representations of constructions* (In The Oxford handbook of construction grammar).

Capirci, O., Cattani, A., Rossini, P., & Volterra, V. (1998). Teaching sign language to hearing children as a possible factor in cognitive enhancement. *The Journal of Deaf Studies and Deaf Education, 3*(2), 135–142.

Cartmill, E. A., Beilock, S., & Goldin-Meadow, S. (2012). A word in the hand: Action, gesture and mental representation in humans and non-human primates. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*(1585), 129–143.

Christiansen, M. H., & Chater, N. (2016). The now-or-never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences, 39.*

Clark, H. H. (2016). Depicting as a method of communication. *Psychological Review, 123* (3), 324.

Cook, S. W., Yip, T. K., & Goldin-Meadow, S. (2012). Gestures, but not meaningless movements, lighten working memory load when explaining math. *Language and cognitive processes, 27*(4), 594–610.

Cormier, K., Smith, S., & Sevcikova-Sehyr, Z. (2015). Rethinking constructed action. *Sign Language & Linguistics, 18*(2), 167–204.

Cowan, N. (2001). Metatheory of storage capacity limits. *Behavioral and Brain Sciences, 24*(1), 154–176.

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science, 19*(1), 51–57.

Croft, W., & Cruse, D. A. (2004). *Cognitive linguistics.* Cambridge University Press.

Dingemanse, M. (2011). Ideophones and the aesthetics of everyday language in a West-African society. *The Senses and Society, 6*(1), 77–85.

Dingemanse, M. (2018). Redrawing the margins of language: Lessons from research on ideophones. *Glossa: a journal of general linguistics, 3*(1). https://doi.org/10.1093/deafed/enn005

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in Cognitive Sciences, 19*(10), 603–615.

Dudis, P. G. (2004). Body partitioning and real-space blends. *Cognitive Linguistics, 15*(2), 223–238.

Elman, J., Bates, E., Johnson, M., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press/Bradford Books.

Emmorey, K. (1993). Processing a dynamic visual—Spatial language: Psycholinguistic studies of American Sign Language. *Journal of Psycholinguistic Research, 22*(2), 153–187.

Emmorey, K. (1995). The confluence of space and language in signed languages. *Linguist Am Sign Lang Introd, 3*, 318–346.

Emmorey, K. (2002). 15 The effects of modality on spatial language: How signers and speakers talk about space. Modality and structure in signed and spoken languages, 405.

Emmorey, K. (2014). Iconicity as structure mapping. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1651), 20130301.

Emmorey, K., & Kosslyn, S. M. (1996). Enhanced image generation abilities in deaf signers: A right hemisphere effect. *Brain and Cognition, 32*(1), 28–44.

Emmorey, K., Kosslyn, S. M., & Bellugi, U. (1993). Visual imagery and visual-spatial language: Enhanced imagery abilities in deaf and hearing ASL signers. *Cognition, 46*(2), 139–181.

Emmorey, K., & Tversky, B. (2002). Spatial perspective choice in ASL. *Sign language & linguistics, 5*(1), 3–26.

Engberg-Pedersen, E. (2003). *From pointing to reference and predication: pointing signs, eyegaze, and head and body orientation in Danish Sign Language* (pp. 269–292). Mahwah, NJ: Pointing: Where Language, Culture, and Cognition Meet. Erlbaum.

Fang, G., Gao, W., & Zhao, D. (2006). Large-vocabulary continuous sign language recognition based on transition-movement models. *IEEE transactions on systems, man, and cybernetics-part a: systems and humans, 37*(1), 1–9.

Fauconnier, G., & Turner, M. (1996). Blending as a central process of grammar. *Conceptual structure, discourse, and language, 113*, 130.

Ferreira, V. S., & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology, 40*(4), 296–340.

Ferrer i Cancho, R. (2004). The Euclidean distance between syntactically linked words. *Physical Review E, 70*, Article 056135.

Ferrer i Cancho, R., & Liu, H. (2014). The risks of mixing dependency lengths from sequences of different length. *Glottotheory, 5*, 143–155.

Fischer, S., & Janis, W. (1990). Verb sandwiches in American Sign Language. *Current trends in European sign language research, 9*, 279–293.

Frick, R. W. (1985). Testing visual short-term memory: Simultaneous versus sequential presentations. *Memory & Cognition, 13*(4), 346–356.

Futrell, R., Mahowald, K., & Gibson, E. (2015). Large-scale evidence of dependency length minimization in 37 languages. *Proceedings of the National Academy of Sciences, 112*(33), 10336–10341.

Geraci, C., Gozzi, M., Papagno, C., & Cecchetto, C. (2008). How grammar can cope with limited short-term memory: Simultaneity and seriality in sign languages. *Cognition, 106*(2), 780–804.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition, 68*(1), 1–76.

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. *Image, language, brain, 2000*, 95–126.

Gibson, E., Futrell, R., Piandadosi, S. T., Dautriche, I., Mahowald, K., Bergen, L., & Levy, R. (2019). How efficiency shapes human language. *Trends in Cognitive Sciences*, 1–19.

Gildea, D., & Temperley, D. (2010). Do grammars minimize dependency length? *Cognitive Science, 34*(2), 286–310.

Goldin-Meadow, S., & Beilock, S. L. (2010). Action's influence on thought: The case of gesture. *Perspectives on Psychological Science, 5*(6), 664–674.

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

Grodner, D., & Gibson, E. (2005). Consequences of the serial nature of linguistic input for sentenial complexity. *Cognitive Science, 29*(2), 261–290.

Haiman, J. (1980). The iconicity of grammar: Isomorphism and motivation. *Language*, 515–540.

Haiman, J. (Ed.). (1985). Iconicity in syntax: proceedings of a Symposium on iconicity in syntax, Stanford, June 24-6, 1983(Vol. 6). John Benjamins Publishing.

Hawkins, J. A. (2004). *Efficiency and complexity in grammars*. Oxford University Press on Demand.

Holler, J., & Wilkin, K. (2011). An experimental investigation of how addressee feedback affects co-speech gestures accompanying speakers' responses. *Journal of Pragmatics, 43*(14), 3522–3536.

Irwin, D. E., & Andrews, R. V. (1996). Integration and accumulation of information across saccadic eye movements. *Attention and performance XVI: Information integration in perception and communication, 16*, 125–155.

Jaeger, T. F. (2006). *Redundancy and syntactic reduction in spontaneous speech*. Stanford University: Doctoral dissertation.

Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology, 61*(1), 23–62.

Jaeger, T. F., & Levy, R. P. (2007). Speakers optimize information density through syntactic reduction. In Advances in neural information processing systems (pp. 849-856).

Jaeger, T. F., & Tily, H. (2011). On language "utility": Processing complexity and communicative efficiency. *Wiley Interdisciplinary Reviews: Cognitive Science, 2*(3), 323–335.

Jiang, J., & Liu, H. (2015). The effects of sentence length on dependency distance, dependency direction and the implications–based on a parallel English–Chinese dependency treebank. *Language Sciences, 50*, 93–104.

Johnston, T., Vermeerbergen, M., Schembri, A., & Leeson, L. (2007). Real data are messy: Considering cross-linguistic analysis of constituent ordering in Auslan, VGT, and ISL. *Visible variation: Comparative studies on sign language structure*, 163–206.

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge University Press.

Kendon, A. (2014). Semiotic diversity in utterance production and the concept of 'language'. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1651), 20130293.

Kendon, A. (2001). Gesture as communication strategy. *SEMIOTICA-LA HAYE THEN BERLIN, 135*(1/4), 191–210.

Kita, S., Van Gijn, I., & Van der Hulst, H. (1997, September). Movement phases in signs and co-speech gestures, and their transcription by human coders. In International Gesture Workshop (pp. 23-35). Springer, Berlin, Heidelberg.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). Package 'lmerTest'. R package version, 2(0).

Labov, W., & Waletzky, J. (1967). J., 1967. Narrative analysis: oral versions of personal experience. Essays on the Verbal and Visual Arts, Seattle, 12, 44.

Lepic, R. (2015). *Motivation in morphology: Lexical patterns in ASL and English*. UC San Diego: Doctoral dissertation.

Lepic, R., & Occhino, C. (2018). A construction morphology approach to sign language analysis. In *In The construction of words (pp. 141–172)*. Cham: Springer.

Levelt, W. J. (1980). On-line processing constraints on the properties of signed and spoken language. In Signed and Spoken Language: Biological constraints on linguistic form(pp. 141-160). Verlag Chemie.

Levelt, W. J. (1981). The speaker's linearization problem. *Phil. Trans. R. Soc. Lond. B, 295*(1077), 305–315.

Liddell, S. K., Vogt-Svendsen, M., & Bergman, B. (2007). A crosslinguistic comparison of buoys. Simultaneity in signed languages: Form and function, 187-215.

Liu, H., Xu, C., & Liang, J. (2017). Dependency distance: A new perspective on syntactic patterns in natural languages. *Physics of Life Reviews, 21*, 171–193.

Liu, H. (2008). Dependency distance as a metric of language comprehension difficulty. *Journal of Cognitive Science, 9*(2), 159–191.

Louwerse, M. M., & Jeuniaux, P. (2010). The linguistic and embodied nature of conceptual processing. *Cognition, 114*(1), 96–104.

Lu, Q., Xu, C., & Liu, H. (2016). Can chunking reduce syntactic complexity of natural languages? *Complexity, 21*(S2), 33–41.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*(6657), 279.

Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods, 49*(4), 1494–1502.

Mahowald, K., Fedorenko, E., Piantadosi, S. T., & Gibson, E. (2013). Info/information theory: Speakers choose shorter words in predictive contexts. *Cognition, 126*(2), 313–318.

McFarland, C. E., & Kellas, G. (1974). Mode of input effects on subject-controlled processes. *Journal of Experimental Psychology, 103*(2), 343.

McNeill, D. (2008). *Gesture and thought*. University of Chicago press.

Metzger, M. (1995). Constructed dialogue and constructed action in American Sign Language. Sociolinguistics in deaf communities, 255-271.

Morgan, G. (2002). The encoding of simultaneity in children's British Sign Language narratives. *Sign Language and Linguistics, 5*(2), 131–165.

Napoli, D. J., & Sutton-Spence, R. (2010). Limitations on simultaneity in sign language. *Language, 86*(3), 647–662.

Napoli, D. J., & Sutton-Spence, R. (2014). Order of the major constituents in sign languages: Implications for all language. *Frontiers in Psychology, 5*, 376.

Occhino, C. (2017). An introduction to embodied cognitive phonology: Claw-5 hand-shape distribution in ASL and Libras. *Complutense Journal of English Studies, 25*, 69.

Oliveira, F. T., & Ivry, R. B. (2008). The representation of action: Insights from bimanual coordination. *Current Directions in Psychological Science, 17*(2), 130–135.

Ortega, G., Sümer, B., & Özyürek, A. (2017). Type of iconicity matters in the vocabulary development of signing children. *Developmental Psychology, 53*(1), 89.

Ortega, G. (2017). Iconicity and sign lexical acquisition: a review. *Frontiers in Psychology, 8*, 1280.

Özyürek, A., & Trabasso, T. (1997). Evaluation during the understanding of narratives. *Discourse Processes, 23*(3), 305–335.

Ozyurek, A. S. L. I., Zwitserlood, I., & Perniss, P. (2010). Locative expressions in signed languages: A view from Turkish Sign Language (TİD). *Linguistics, 48*(5), 1111–1145.

Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory & Cognition, 17*(4), 398–422.

Perniss, P., Thompson, R., & Vigliocco, G. (2010). Iconicity as a general property of language: Evidence from spoken and signed languages. *Frontiers in Psychology, 1*, 227.

Perniss, P., & Vigliocco, G. (2014). The bridge of iconicity: From a world of experience to the experience of language. Phil. Trans. R. Soc. B, 369(1651), 20130300.

Perniss, P., Zwitserlood, I., & Özyürek, A. (2015). Does space structure spatial language?: A comparison of spatial expression across sign languages. *Language, 91*(3), 611–641.

Perniss, P. M. (2007a). *Space and iconicity in German sign language (DGS)*. Radboud University Nijmegen Nijmegen: Doctoral dissertation.

Perniss, P. M. (2007b). Achieving spatial coherence in German sign language narratives: The use of classifiers and perspective. *Lingua, 117*(7), 1315–1338.

Piantadosi, S. T., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences, 108*(9), 3526–3529.

Pietrandrea, P., & Russo, T. (2007). Diagrammatic and imagic hypoicons in signed and verbal languages. *Empirical Approaches to Language Typology, 36*, 35.

Pizzuto, E. A., Rossini, P., Sallandre, M. A., & Wilkinson, E. (2006, December). Deixis, anaphora and highly iconic structures: Cross-linguistic evidence on American (ASL), French (LSF) and Italian (LIS) Signed Languages. In TISLR9(pp. 475-495). Editora Arara Azul, Brazil.

Antinoro Pizzuto, Elena, Paolo Rossini, Marie-Anne Sallandre and Erin Wilkinson 2008. Deixis, anaphora and highly iconic structures: Cross-linguistic evidence on American (ASL), French (LSF) and Italian (LIS) Signed Languages. In: Ronice Müller de Quadros (ed.), Proceedings of TISLR9, Theoretical Issues in Sign Language Research Conference, 475–495. Petrópolis, RJ: Editora Arara Azul.

Race, D. S., & MacDonald, M. C. (2003). The use of "that" in the production and comprehension of object relative clauses. In Proceedings of the Annual Meeting of the Cognitive Science Society (Vol. 25, no. 25).

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences, 20* (9), 676–688.

Risler, A. (2007). A cognitive linguistic view of simultaneity in process signs in French Sign Language. *AMSTERDAM STUDIES IN THE THEORY AND HISTORY OF LINGUISTIC SCIENCE SERIES, 4*(281), 73.

Rogers, K. L. (2012). American Sign Language verb categories in constructed action.

Rudner, M., & Rönnberg, J. (2008). Explicit processing demands reveal language modality-specific organization of working memory. *Journal of Deaf Studies and Deaf Education, 13*(4), 466–484.

Russo, T. (2004). Iconicity and productivity in sign language discourse: An analysis of three LIS discourse registers. *Sign Language Studies, 4*(2), 164–197.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied, 74*(11), 1.

Sumer, B., Zwitserlood, I., Perniss, P. M., & Ozyurek, A. (2013). Acquisition of locative expressions in children learning Turkish Sign Language (TİD) and Turkish. In Current directions in Turkish Sign Language research (pp. 243-272). Cambridge Scholars Publishing.

Supalla, S., Siple, P., & Fischer, S. (1991). Manually coded English: The modality question in signed language development. *Theoretical issues in sign language research, 2*, 85–109.

Sutton-Spence, R., & Braem, P. B. (2013). Comparing the products and the processes of creating sign language poetry and pantomimic improvisations. *Journal of Nonverbal Behavior, 37*(4), 245–280.

Tagliamonte, S., Smith, J., & Lawrence, H. (2005). No taming the vernacular! Insights from the relatives in northern Britain. *Language Variation and Change, 17*(1), 75–112.

Taub, S. F. (2001). *Language from the body: Iconicity and metaphor in American Sign Language*.

Temperley, D. (2007). Minimization of dependency length in written English. *Cognition, 105*(2), 300–333.

Temperley, D., & Gildea, D. (2018). Minimizing syntactic dependency lengths: Typological/cognitive universal? *Annual Review of Linguistics, 4*, 67–80.

Ungerer, F. (1999). Diagrammatic iconicity in word-formation. *Form miming meaning: Iconicity in language and literature*, 307–324.

Vermeerbergen, M., Leeson, L., & Crasborn, O. A. (Eds.). (2007). Simultaneity in signed languages: Form and function(Vol. 281). John Benjamins Publishing.

Vinson, D., Thompson, R. L., Skinner, R., & Vigliocco, G. (2015). A faster path between meaning and form? Iconicity facilitates sign recognition and production in British sign language. *Journal of Memory and Language, 82*, 56–85.

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 27*(1), 92.

Wilcox, S. (2004). Cognitive iconicity: Conceptual spaces, meaning, and gesture in signed language. *Cognitive linguistics, 15*(2), 119–148.

Wilson, D., & Wharton, T. (2006). Relevance and prosody. *Journal of Pragmatics, 38*(10), 1559–1579.

Wilson, M., & Emmorey, K. (1997). Working memory for sign language: A window into the architecture of the working memory system. Journal of Deaf Studies and Deaf Education, 121-130.

Wilson, M., & Emmorey, K. (2006). No difference in short-term memory span between sign and speech. *PSYCHOLOGICAL SCIENCE-CAMBRIDGE, 17*(12), 1093.

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: A professional framework for multimodality research. In 5th International Conference on Language Resources and Evaluation (LREC 2006) (pp. 1556-1559).

Woodman, G. F., Vecera, S. P., & Luck, S. J. (2003). Perceptual organization influences visual working memory. *Psychonomic Bulletin & Review, 10*(1), 80–87.

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science, 13*(2), 168–171.

Zwaan, R. A., & Yaxley, R. H. (2003). Spatial iconicity affects semantic relatedness judgments. *Psychonomic Bulletin & Review, 10*(4), 954–958.