

MANON LELANDAIS
UNIVERSITY OF NANTES

GAËLLE FERRÉ
UNIVERSITY OF NANTES

THE VERBAL, VOCAL, AND GESTURAL EXPRESSION OF (IN)DEPENDENCY IN TWO TYPES OF SUBORDINATE CONSTRUCTIONS

CITATION

Lelandais, M. & Ferré, G. (2019). The verbal, vocal, and gestural expression of (in)dependency in two types of subordinate constructions. *Journal of Corpora and Discourse Studies*, 2:117–143

ABSTRACT

Based on a video recording of conversational British English, this paper tests within the framework of Multimodal Discourse Analysis whether two different subordinate structures are integrated to the same degree in their environment. Subordinate constructions have been described in linguistics as dependent forms elaborating on primary elements of discourse. Although their verbal characteristics have been analysed in depth, few studies have focused on the articulation of the different communicative modalities in their production or provided a qualified picture of their integration. The main hypothesis is based on the capacity of subordinate constructions to show distinct forms of autonomy depending on their syntactic type, thus expressing different degrees of break. Beyond showing that subordinate constructions are not dependent to the same degree depending on how speakers use the prosodic and kinetic modalities to express greater (in)dependency, the results suggest that the creation of a break mainly relies on prosodic cues. Changes in the modal configuration throughout the sequence suggest modalities are dynamic and flexible resources for integrating or demarcating subordinate constructions in function of their syntactic type.

KEYWORDS

syntax; boundaries; subordination;
prosody; gesture; multimodal discourse
analysis

CONTACT

Manon Lelandais, Département du Monde Anglophone, UFR Langues Littératures
Cultures et Sociétés Etrangères, Université Paris 3 Sorbonne Nouvelle. 5 Rue de l'École
de Médecine, 75006, Paris. manon.lelandais@sorbonne-nouvelle.fr

DOI

10.18573/jcads.4

ORCID

0000-0003-4962-0334

ISSUE DOI

10.18573/jcads.v2

LICENSE

© The authors. Available under the terms of the CC-BY 4.0 license

The verbal, vocal, and gestural expression of (in)dependency in two types of subordinate constructions

Manon Lelandais

University of Nantes

Gaëlle Ferré

University of Nantes

1. Introduction

This article discusses subordination in spontaneous speech, more specifically the sequences containing subordinate constructions operating at the syntactic level of modification (e.g., Huddleston & Pullum, 2002).

In syntactic studies, *modifiers* refer to elements specifying or elaborating upon some primary features (Halliday, 1985), often described as additions to propositional contents in the host or embedding structure (Quirk *et al.*, 1985, p. 1058; Huddleston & Pullum, 2002, p. 1048). Two semantic types are distinguished when describing dependency relations (van Rijn, 2017). While some heads inherently require reference to a dependent, which may therefore be considered the head's argument, other dependents are not inherently presupposed by their head and are considered modifiers. They merely provide a further semantic characterisation of the referent (or state of affairs) expressed by the head, or supplement the head with additional information. Well-known examples of modifiers are relative clauses (van Rijn, 2017).

Instead of comparing subordinate clauses to non-subordinate clauses in discourse, the study aims to identify differences between two syntactic types of subordinate constructions in terms of demarcation, through their multimodal expression. We propose that subordinate clauses, and especially the category of modifiers, are not as unified and dependent as described, and that different degrees of verbal, vocal, and gestural break can be distinguished according to their syntactic type. In the research presented here, the subordinate constructions (SC) under study encompass the two most widespread types of finite clauses functioning as modifiers in our oral corpus of spontaneous interaction (which is described in Section 3. Corpus and Methodology of the paper): restrictive relative clauses (RRCs henceforth) and appositive relative clauses (ARCs henceforth), as illustrated in Excerpts 1–2.

A restrictive relative clause modifies a nominal expression. A relation of co-referentiality holds between the nominal referent and some participant in the process designated by the relative clause. This participant, or *pivot*, has a semantic role in both the relative clause and the matrix clause containing the modified nominal (Langacker, 2008, p. 424). According to Langacker (1991), the meaning of a common noun is a mere type of entity. The designation of instances requires a full NP, in which the *type specifications* conveyed by the common noun are tied to the speech exchange by determiners. A restrictive relat-

ive clause ‘restricts the head noun’s type specification’ (Langacker, 1991, p. 432), i.e. it delineates a subtype of the general type of entity designated by the head noun. In Excerpt 1, the restrictive relative clause increases the relevance of ‘the reasons’, creating a subcategory for this referent.

(1) Restrictive relative clause

Tim	L	the reasons
	Sc	<i>they gave #</i>
	R	i mean Blair was also #

This paper focuses on RRCs introduced by \emptyset , *that*, and *where* as relative pronouns. They have been analysed as mainly working towards defining the antecedent (Cotte, 2008). Unlike *which*, which links the referent it modifies with the contextual situation (Melis, 2008), these relative pronouns have been described as object-oriented (i.e., defining a property of the object). This construction allows speakers to provide the co-speaker with more complex information about the antecedent than in non-relative structures, without the co-speaker having trouble processing it. The antecedent opens an informational frame about the referent it describes, while the relative pronoun indicates that the informational frame about this referent is about to be completed (Muller, 2006).

Although also introduced with a relative pronoun, ARCs are not invoked to single out a nominal referent, but to make an additional comment about it (Langacker, 2008, p. 429). Their modifying scope varies from a single nominal referent to a verb phrase or a whole clause (Longacre, 1985). In Excerpt 2, the appositive relative clause qualitatively evaluates ‘black pudding’, which can however be identified independently.

(2) Restrictive relative clause

Tom	L	but then again
		i’ll happily eat black pudding
	Sc	<i>which i know is disgusting #</i>
	R	blood and guts and

This study focuses on ARCs introduced by *which* as a conjunction. Unlike *that*, *which* introduces a subjective comment bearing on the relationship between two propositions. In Excerpt 2, Tom links ‘black pudding’ with the non-neutral adjective ‘disgusting’. Setting up a functional distinction between several types of ARCs, Melis (2008) evokes comment appositive clauses (as in Excerpt 2 above) and continuation appositive clauses. In the latter type, the relative pronoun only intervenes as an inter-propositional relator. The literature generally agrees on the fact that ARCs show several characteristics that are typically

associated with non-subordinate clauses (Krifka, 2007)¹. Their capacity to form distinct illocutionary acts (e.g., Peterson, 1999; Holler, 2005) is one such property.

Subordinate constructions are generally defined as dependent on another predication from a micro- or macro-syntactic point of view (Tomlin, 1985; Lehmann, 1988, among others). Lehmann (1988, p. 181) proposes a definition of the notion of dependency, linked to that of integration: '[a] grammatical relation R connecting syntagms X and Y is a relation of dependency if X occupies a grammatical slot of Y or vice versa. In a dependency relation, Y depends on X if X determines the grammatical category of the complex and thus its external relations.' In sum, a clause B is dependent on a clause A if B lacks functional autonomy, and is linked to one of A's components: B is hierarchically down-graded in its link to A. In Cognitive Grammar, dependency is explained in terms of conceptualisation. According to Langacker (2008, p. 199), certain structures 'require the support of others [...] for their own manifestation. As a consequence, a dependent structure refers schematically to an autonomous, supporting structure as an intrinsic aspect of its own characterization.' A modifier, for instance, is dependent on an autonomous 'head' (p. 203). Dependency is an asymmetrical relation between two connected structures, with one structure overriding the other. For instance, in 'the flowers I bought today are for my sister', the referential scope of the main clause overrides that of the relative clause 'Ø I bought today'. The structural centre of the main clause (i.e. the predicate 'are for my sister') also dominates the predicate in the subordinate clause, as the predicate in the subordinate clause elaborates upon a component of the main clause. The complex structure is a layered organisation in which one clausal relationship incorporates the other as a participant.

In this study bearing on oral spontaneous speech, we consider a structure B to be dependent on a structure A if B is integrated in A, in the sense that there is no boundary mark between A and B. By contrast, a structure A is considered to be autonomous if A features boundary marks. The different boundary marks taken into account to determine dependency in our corpus are fully described in Section 2. Theoretical Background.

While subordinate constructions are broadly defined as dependent, the literature shows little consensus in defining clear scopes and boundaries for these structures. While they are seen as embedded elements (Jackendoff, 1977), Jespersen (1927), Fabb (1990), and Peterson (1999) consider some relative constructions as exterior to the syntactic structure of the main clause. More specifically, the differentiation between RRCs and ARCs on syntactic grounds is problematic (Borsley, 1992; Arnold & Borsley, 2008).

1 While the functional distinction between continuation ARCs (moving the discourse on and contributing to the foreground) and comment ARCs (bringing background information in discourse) is widely accepted (Lambrecht, 1996; Holler, 2005), a recent study (Lytvynova & Dao, 2014) calls it into question, asserting that none of these two categories can be likened to autonomous discourse units, from both grammatical and pragmatic points of view. Potts (2005) also describes the content of ARCs as non-asserted, as it cannot be directly questioned.

While appositive relatives are derived from coordination for Burton-Roberts (1999) and De Vries (2006), Quirk *et al.* (1985, p. 1257) and Biber *et al.* (1999, p. 135) call for special levels of representation in subordination, with ‘telescoped relatives’ and ‘peripheral elements’ respectively. Appositive relatives are also classified as adverbials (Biber *et al.*, 1999, p. 853). Finally, in the rich framework of syntactical relations proposed by Matthiessen & Thompson (1988, p. 238), adverbials and non-restrictive relatives are ‘less subordinate’ than other structures, belonging to relations of hypotactic combination instantiating hypotaxis, where dominant and dependent clause are syntactic sisters, rather than embedding, where the subordinate clause is a constituent of the dominant clause. This study therefore questions whether these constructions all express the same degree of dependence (understood as integration, i.e., lack of boundary marks) upon their co-text.

If some substantial work has focused on the relation of subordinate clauses to their ‘hosts’ from the perspectives of syntax (Haiman & Thompson, 1984; De Vries, 2006; Haegeman, 2010) or prosody (Bolinger, 1984; Couper-Kuhlen, 1986, among others), the development of analytical tools and schemes now facilitates an account of subordinate constructions as multimodal phenomena.

In face-to-face conversation, participants negotiate meaning through multimodal contributions, in which the linguistic resources of speech interface with some kinetic modes, such as gaze direction, eyebrow and head movements, as well as hand gestures (although posture and facial expressions can also be used by speakers in a communicative way, they were not considered in the present work as we did not want to multiply the number of parameters used in the study). These modes do not work independently of the others, although a particular mode may weigh more than the others at some points (Norris, 2004). Within speech itself, the vocal or the verbal mode may stand out as more prominent at particular points in time.

Viewing language as integrating speech and gesture in an organised system enables investigation of the production process of subordinate constructions, focusing on boundaries (i.e., markers of demarcation in terms of independence) on several levels. The main hypothesis is based on the capacity of these constructions to show distinct forms of autonomy depending on their syntactic type. Different degrees of break are consequently identified from the new perspective afforded by multimodality, providing a qualified picture of their multimodal insertion in discourse.

The study of subordination in spontaneous speech from a multimodal point of view gives new perspectives on discourse production, particularly on the flexibility of discourse planning and modelling. More information on real-time discourse production with a particular focus on boundaries benefits such areas of study as natural language processing and language acquisition. Online speech processing can drive learning and inform pedagogy (Gilbert *et al.*, 2016).

The first part of this paper presents a review of the literature concerned with the notion of subordination, focusing on its syntactic, prosodic and gestural acceptations. A detailed description of our corpus and methodology ensues, immediately followed by analysis and discussion of the data.

2. Theoretical Background

2.1. Syntax

In the categorial division of clause complexes into a classification that comprises two uneven and complementary subgroups, i.e., a main clause and a subordinate, modifiers are viewed as 'optional' constituents functioning at a phrasal or clausal level. This classification arises from the concept of minimal utterances: some elements of the message are deemed semantically necessary without standing as constitutive elements. However, this semantic necessity has been queried by a number of linguists (e.g., Haiman & Thompson, 1984; Chafe, 1988; Smessaert *et al.*, 2005), and described as imprecise for analysing spontaneous speech, especially regarding the nature of introductory elements.

Because semantic necessity is felt to be imprecise, other criteria are suggested to evaluate clausal combination, in a hierarchy of syntactic and semantic relations: a close semantic relation between two clauses correlates to a tight syntactic linkage (van Valin, 1984). Clauses are units comprising an essential nucleus (containing the predicate that corresponds to an event, process or state, and its core complements), and an optional periphery (corresponding to the spatiotemporal frame such as localisation or environment; van Valin, 1984; Halliday, 1985). A clause attached to the nuclear components of another clause demonstrates a stronger bond than a clause linked to the peripheral elements of another.

These criteria all encourage investigation of clause linkage relying on a wider, more detailed set of syntactic and semantic parameters (van Valin, 1984), or going beyond the micro-syntactic frame in observing not only governing relations, but also modal and illocutionary relations (Thompson, 2002; Heringa, 2007).

If very little work has been conducted on the subordinate forms under study from a multimodal perspective, a large body of research has detailed prosodic subordination (Bolinger, 1984; Local, 2007; Wells, 2006, among others), while other studies on co-verbal gestures have exposed visual means of subordination (Enfield, 2009; Streeck, 2009), sometimes but not necessarily correlated to verbal subordination.

2.2. Prosody

Prosodic subordination is essentially achieved through intonation (Bolinger, 1984). Throughout a vocal paragraph, pitch height naturally declines in a progressive manner. A subordinate unit is signalled by downward changes in key (i.e., major levels in a speaker's

pitch range) or in pitch height (Lelandais & Ferré, 2016). Intonation can convey subordinating information that is not marked with verbal means.

To integrate a prosodic unit into an adjacent segment, *fundamental frequency* (F0), which participates in the perception of pitch, generally rises on the final syllable of the inserted segment, indexing this segment as prefacing further speech, continuing the paragraph and the point being treated (Wells, 2006). A downstepped tone compared to a preceding high tone corresponds to the general neutral relationship between two prosodic groups, often used to express seamless continuity.

By contrast, a variation on the initial syllable signals a boundary. Likewise, a discourse segment featuring a low final syllable (termination contour) does not embed the following segment, and is autonomous regarding what follows. Boundaries in speech can also be created with a variation of tempo (Crystal, 1969). Silent pauses conjointly participate in the segmentation of discourse (Local, 1992).

2.3. Gestures

The important role of co-speech gestures in linguistic production has been shown in pragmatics (Lascarides & Stone, 2009), cognitive linguistics (Sweetser, 2006), and psycholinguistics (McNeill, 2005).

2.3.1. Hand gestures

Some studies (Cassel & McNeill, 1990; Calbris, 2011) have shown that some gesture features participate in the maintenance of coherence and cohesion (Halliday & Hasan, 1976) in speech. Likewise, representing referents through hand gestures is a cumulative process, often achieved through a series of several gesture units. Reference is maintained through cohesiveness of space, handedness, and/or form, including style of movement (McNeill & Levy, 1993; Streeck, 2009).

Two speech segments can also be related through their production in co-occurrence with a single gesture unit (Enfield, 2009). On the contrary, a rest position for both hands can signal a boundary in discourse (Calbris, 2011), contrasting with a preceding gesture sequence. Likewise, hand beat gestures can single out particular entities (Cavé *et al.*, 1996). These are connected to discourse structure in their function (Kendon, 1972; De Kok & Heylen, 2009).

2.3.2. Gaze direction

Gaze often moves away from the co-speaker for discourse elaboration as soon as the speaking turn is taken and secured (Beattie, 1978; De Kok & Heylen, 2009). A change in gaze direction towards the co-speaker announces a discourse boundary (Beattie, 1978; De Kok & Heylen, 2009) or an appeal to the co-speaker (Goodwin, 1979; Holler *et al.*, 2014). When a discourse sequence is successfully completed, its participants routinely turn their

gaze away from one another, whereas one or both parties maintain gaze on the other when the sequence needs an expansion (Rossano, 2012).

2.3.3. Head and eyebrow movement

Head gestures, particularly head beats (i.e., brief downward chin movements) and nods, mark out the rhythmic organisation of the utterance (Cavé *et al.*, 1996). Eyebrow movement, especially rises, may also provide segmental information (Granström & House, 2005).

3. Corpus and Methodology

3.1. Corpus recording

The corpus used for this study, ENVID (Lelandais & Ferré, 2016), is a collection of dialogues in British English. This collaborative corpus gathers videos recorded between 2000 and 2012. Five dialogues were selected, making up a total of 2 hours and 10 minutes of interaction. Each interaction was recorded in a soundproof studio, guaranteeing its prosodic treatment. The participants are British people aged 20 to 23. Each participant had a lavalier microphone, which provided two separate audio tracks. Two audio files corresponding to each microphone were created in a WAV format, so as to facilitate the analysis of overlapping speech. The native video recordings were transformed into MPEG-4 stereo files, with a rate of 25 frames per second. Each dialogue had a single MPEG-4 file, juxtaposing the images of both cameras for the interactions filmed with a camera facing each participant.

Each participant is filmed in a static, wide-angled shot, facing or three-quarters turned towards their interlocutor. They are visible at least from head to chest, the cameras capturing subtle face movements but also rendering an overview of their upper body parts. In two of the dialogues, a single fixed camera faces the two participants. In the other three interactions, a fixed camera faces each participant.

3.2. Corpus transcription

The corpus was first transcribed in Praat (Boersma & Weenink, 2013) by the two authors adopting a standard orthographic spelling. Subordinate constructions were then localised and coded on a separate track as SC. All the annotations made in Praat were afterwards exported into Elan (Sloetjes & Wittenburg, 2008), a video annotation tool, in order to relate information in the different domains.

3.2.1. Syntactic annotation

A total of 141 syntactic constructions were annotated in the corpus by one of the authors, representing 2.82% of the total speaking time (i.e., 1.19 form/min). Among these subordinate constructions were 83 restrictive relative clauses (1.65% of speaking time, 0.7 form/min) and 58 ARCs (1.17% of speaking time, 0.5 form/min).

Forty occurrences of each syntactic type (ARCs and restrictive relative clauses) were selected for a balanced comparison, making up a total of 80 forms. The selection targeted occurrences without an interruption, surrounded with immediate left and right co-texts other than a single silent pause yielding the speaking turn. We also made sure that our selection of syntactic constructions was balanced across speakers. This was significantly important for the analysis of gestures, so as to avoid any bias due to inter-speaker gestural variability.

The selected occurrences were classified according to their syntactic type in Praat (restrictive relative clause, appositive relative clause). A second track delimitates the environment of these clauses: the preceding tone-unit or part of tone-unit was labelled L (left co-text), the subsequent one labelled R (right co-text).

In order to establish reliability of the clause type classification (restrictive relative clause, adverbial clause, appositive relative clause), a second expert coder judged 20% of the data that had been classified by the original coder. The agreement between coders was 100%.

3.2.2. Prosodic coding

The corpus was segmented into tone-units, according to the British school of intonation (Crystal, 1969; Wells, 2006) based on dynamic pitch contours. Tone-units correspond to Intermediate Phrases in the ToBI system (Beckman *et al.*, 2005).

Although our annotation relies on a different theoretical framework from that of the ToBI system, our interest in boundaries can be drawn close to the systematic annotation of break indices in the ToBI line of work (e.g., Beckman *et al.*, 2005).

The Momel-Intsint algorithm (Hirst, 2007; Bigi, 2012) was used for the automatic annotation of the F0 target points in the signal. Annotations are made in two respects: the algorithm notes pitch height (in Hz) on target syllables, which allows us to calculate mean F0 values for specific segments. The algorithm also codes symbolic (relative) values of intonation, in which each measured F0 value is compared to preceding ones, i.e. significant changes in the F0 curve either regarding the speaker's pitch range (Top, Bottom) or regarding the neighbouring tones or sequences of tones (Upstep, Downstep, Same, Low, High). We are here particularly interested in values which indicate a significant pitch reset (Top, Bottom), or a significant change in pitch key (Upstep – change towards higher pitch range, Downstep – towards lower pitch range). We are also interested in the value Same which, if found in greater number in our sequences, would indicate that there is no break between the different elements of the sequence.

Within each segment of the sequences under study, the nature of each nuclear contour (fall; fall-rise; rise; rise-fall; flat) was also coded manually by one of the authors. Pitch key was then annotated in regards to each speaker's specific range (high; mid; low) on both the whole segments (L, Sc, R) and the boundary (initial and final) syllables in these segments.

In order to establish reliability of the nuclear contour classification, a second expert coder judged 20% of the data that had been classified by the original coder. The agreement between coders was 81.9%.

3.2.3. Gesture coding

Communicative gestures were coded in Elan (Sloetjes & Wittenburg, 2008), in which hand gestures, head and eyebrow movement as well as gaze direction were manually coded by the same two expert coders, following the parameters proposed by Bressemer & Ladewig (2011).

Gesture annotation was based on gesture phrases (Kendon, 2004). Each gesture phrase was considered to start at the onset of the gesture and to end at the return to rest position if there was one. In the case of two consecutive gestures, the first gesture phrase ends at a significant change in shape and/or trajectory.

Head movements were labelled into nods, shakes, tilts, beats, or jerks. In separate tracks, gaze direction was annotated as either towards the co-participant or away, eyebrow movement distinguished between rise and frown, and hand gestures were categorised into iconics, metaphoric, pointings, beats, emblems, butterworths, and adaptators, drawing mainly from McNeill's typology (2005). As hand gestures may have several dimensions, two values could be noted and counted if need be.

Hand gestures were coded considering their link with co-occurring speech and their relationship to lexical affiliates (Kipp *et al.*, 2007). Ambiguous types were resolved with discussion between the two coders and agreement was reached on the main dimension of gesture types. Iconics are 'images of concrete entities and/or action', whereas metaphoric are 'images of the abstract' involving a metaphoric use of form and/or space (McNeill, 2005, p. 39). Pointing gestures are deictics whereas beats are linked with speech rhythm (McNeill, 1992, p. 80), emblems are conventionalised signs and butterworths are disorganised gestures made in lexical retrieval. Adaptators, i.e., self-contact gestures used for comfort like scratching one's head, were included given their high number and the fact that they give the experimenters information on the organisation of turns, being more frequent when the participant is listening.

In order to establish reliability of the gesture type classification, a second expert coder judged 20% of the data that had been classified by the original coder. The agreement between coders was 100% for gaze direction, 96.4% for eyebrow movement, 81.3% for head movement, and 72.1% for hand gestures.

3.2.4. Working hypotheses

Based on the theoretical background defined by the literature, a specific list of syntactic, rhythmical, intonational, and visual cues is taken into account to survey different types of boundaries. If the constructions are not autonomous, they should be syntactically governed (i.e., they should fall under the scope of the main predicate, which determines their

syntactic position in the segment) and/or show restricted modal and semantic autonomies, as these features are associated with subordinate segments (van Valin, 1984).

At a prosodic level, embedded constructions are expected to be integrated into the same tone-unit as the main structure (Bolinger, 1984), or to show continuation contours (final rise conveying that the speaker will hold the floor; Hirschberg & Grosz, 1992). Subordinate constructions should be uttered in a low or mid-key, the usual declination line of the paragraph being followed without any break (Wennerstrom, 2001). Subordinate constructions should not cause any important change in rhythm, featuring few pauses (Local, 1992).

As far as the co-verbal gestures are concerned, non-autonomous constructions are expected to be produced with gestural cohesive linkages such as the use of the same gesture unit as their sequential environment (Enfield, 2009). No hand beat is expected around subordinate structures. If subordinate constructions are integrated in their co-text, they should be realised without any change in gaze direction towards the co-speaker (Beattie, 1979; De Kok & Heylen, 2009), and without any eyebrow rise (Cavé *et al.*, 1996).

4. Results

The analysis evaluates the autonomy of subordinate constructions on a linear organisational mode (integration vs. demarcation). We test whether these constructions mainly create a break or whether they are preferentially integrated in their left and/or right co-text. After identifying and measuring the most relevant cues expressing boundaries in the different modalities drawing on our assumptions, the two syntactic types can be differentiated in terms of autonomy. Restrictive relative clauses are more integrated than ARCs. This section presents raw results. Examples will be provided in relation with these results in Section 5. Discussion, in which the particularities are detailed for the two syntactic types. The subsequent series of tables (Table 1, Table 2, Table 3) present the cues taken into account, along with their distribution in the two syntactic types of subordinate constructions. Our analysis is selective in choosing which elements to comment upon for reasons of space, and highlights different cues for each clause type. However, each table gives the number of occurrences (out of 40) featuring the cue taken into account, and the percentage this number represents. Statistical F-tests were run instead of ANOVAs, as the sets of data do not follow a normal distribution and are relatively small. These tests aim to detect a significant difference between the two different types of subordinate constructions (ARCs and RRCs). An asterisk (*) signals a statistically significant result ($p \leq .050$)².

2 In this paper, any differences between two compared groups will be seen as statistically significant if $p \leq .050$; that is, the null hypothesis is rejected if $p \leq .050$.

Table 1 describes the macro-syntactic boundary features that are considered in our analysis. For reasons of space and clarity, this paper focuses on the verbal boundary cues at a macro-syntactic level only: boundaries at the level of discourse are not included.

Macro-syntactic feature	ARCs		RRCs	
	<i>f</i>	%	<i>f</i>	%
Total government	4	10*	30	75
Syntactic government; illocutionary autonomy	5	13	7	18
Syntactic autonomy; illocutionary government	1	3	1	3
Total autonomy	30	75*	2	5

Table 1: Occurrence of the macro-syntactic features taken into account to determine the dependency or autonomy (grey zones) of each syntactic type. Figures are given out of the total of 40 constructions.

In the macro-syntactic part, we focus on the relation between the subordinate construction and other verb phrases in the sequence (located in L, the preceding tone-unit, or in R, the following tone-unit). Following van Valin (1984), Thompson (2002), and Heringa (2007), we identify degrees of integration, in which modal and pragmatic parameters intervene. Total government describes integration, in which another verb phrase in the sequence determines the syntactic position of the subordinate clause and imposes functional restrictions, such as tense, mood, negation, agreement, and/or position in the syntactic unit. Total government also means that the subordinate construction is not pragmatically free (i.e., its truth value as a speech act cannot be separated from that of L or R). We distinguish two intermediate configurations, in which the subordinate construction is either syntactically or pragmatically autonomous. Total autonomy represents the cases in which the subordinate construction is not determined by another verb phrase in any syntactic means, and forms a distinct illocutionary unit (i.e., a speech act).

Table 1 shows that only 10% of ARCs are characterised with total government. This is significantly different from RRCs ($F(39, 39) = 2.08$ $p = .012$; independent variable: syntactic type, dependent variable: total macro-syntactic government). Likewise, ARCs feature significantly more occurrences showing total macro-syntactic autonomy (75% of occurrences; $F(39, 39) = 3.95$ $p < .001$; independent variable: syntactic type, dependent variable: total macro-syntactic autonomy). Restrictive relative clauses are more characterised with total government, with 75% of governed occurrences.

Table 2 features the prosodic boundary cues we have included in the analysis.

Prosodic feature	ARCs		RRCs	
	<i>f</i>	%	<i>f</i>	%
Shorter duration	13	32.5	20	50*
Extra-constituent silent pauses (% pausing time)	43	56.4*	19	24.1
Intra-constituent silent pauses (% pausing time)	20	42	17	23
INTSINT pitch variation values	99		51	
Initial pitch upstep (boundary tones)	10	25	6	15
SC features a distinct tone-unit from that of L and from that of R	38	95*	22	45
No continuation contour on L or SC	31	77*	3	7.5
Significant pitch gap (>20Hz) in SC	40% of speakers		10% of speakers	

Table 2: Occurrence of the prosodic features taken into account to determine the dependency or autonomy (grey zones) of each syntactic type. Figures are given out of the total of 40 constructions.

As far as rhythm is concerned, changes in duration, as part of tempo, are strong indicators of a boundary in speech (Crystal, 1969). Pauses segment the flow of speech (Local, 1992); we have considered both extra-constituent pauses (i.e., pauses separating two different tone-units) and intra-constituent pauses (i.e., pauses inside tone-units). Filled pauses are not included in this analysis, as disfluencies would have been treated as verbal phenomena at the level of discourse, and as audible inbreath is a source of important variation among the speakers of our corpus. As far as intonation is concerned, we focus on INTSINT's automatic demarcative pitch values, as well as on significant gaps in pitch height between the subordinate clause and the preceding segment (Wennerstrom, 2001). The occurrences in which the subordinate construction features its own tone-unit (i.e. is realized in a distinct tone-unit from that of L and from that of R) are taken into account. We also measure the number of continuation contours (both on the subordinate clause and the preceding segment), which indicate a strong link with the following tone-unit (Hirschberg & Grosz, 1992).

Table 2 shows that ARCs feature significantly more extra-constituent silent pauses than restrictive relatives (56.4% of pausing time; $F(42, 18) = 2.22, p = .035$; independent variable: syntactic type, dependent variable: duration of extra-constituent silent pauses in seconds). Appositive relative clauses are not only significantly realised more often with distinct tone-units than restrictive relatives (with 95% of occurrences featuring their own tone-unit; $F(39, 39) = 0.19, p < .001$; independent variable: syntactic type, dependent variable: distinct tone-unit on SC), but also show significantly less continuation contours than restrictive relatives ($F(39, 39) = 2.51, p = .002$; independent variable: syntactic type, dependent variable: number of intonational continuation contours), in that 77% of the sequences containing them do not feature any.

Restrictive relative clauses feature only one disruptive prosodic cue. They show a distinct duration pattern, in that SC is the shortest segment of the sequence in which it is inscribed (L: $F(39, 39) = 4.03, p < .001$; R: $F(39, 39) = 2.15, p = .009$; independent variable: sequence segment (L, SC, R), dependent variable: speech segment production duration in seconds). However, the difference with ARCs is not significant.

Table 3 shows the gestural parameters we have considered as boundary markers.

Gestural feature	ARCs		RRCs	
	<i>f</i>	%	<i>f</i>	%
Overlapping hand gestures	14	19.7*	34	47.9
(L-SC	8	11.3	16	22.5
SC-R	4	5.6	8	11.3
L-R)	2	2.8	10	14.1
Hand beats (% total hand beats)	7	23.3	16	53*
Hand beats (% total head beats)	19	42	9	20
No change in gaze direction throughout L-SC-R	10	25	27	67.5
Eyebrow rises (% total rises)	21	63.6*	5	15.2

Table 3: Occurrence of the gestural features taken into account to determine the dependency or autonomy (grey zones) of each syntactic type. Figures are given out of the total of 40 constructions.

We measure the number of gestures that are produced in overlap with two or three different tone-units during the sequences under study, specifying the tone-units in question (overlap from L to SC, from SC to R, or from L to R). The overlaps in question include

the subordinate construction in the same gesture unit as the co-text (Enfield, 2009; Streeck, 2009), showing no boundary. The subordinate type produced with the smallest proportion of gestural overlap is thus taken into account. Beat gestures (hand and head gestures alike) occurring in isolation (i.e., the beats that are not included in a catchment — a recurrent pattern of beats providing cohesion in a specific discourse sequence) are also included for their segmental properties (Kendon, 1972; Cavé et al., 1996). Changes in gaze direction patterns (De Kok & Heylen, 2009) are investigated, as well as the number of eyebrow rises (Granström & House, 2005).

Table 3 shows that ARCs are uttered together with distinct gesture units: the distribution of overlapping hand gestures between L, SC, and R is significantly lower than for restrictive relatives (19.7% of occurrences, in which only 2.8% feature gestures overlapping from L to R; $F(39, 39) = 3.95, p < .001$; independent variable: syntactic type, dependent variable: number of overlapping hand gestures). Eyebrow rises also characteristically feature in ARCs, as not only do they occur significantly more in SC than in the preceding segment (L: $F(39, 39) = 1.72, p = .046$; independent variable: sequence segment (L, SC, R), dependent variable: number of eyebrow rises), but they also occur significantly more often with this syntactic type than in the other (21% of ARCs are produced with eyebrow rises; $F(39, 39) = 2.28, p = .005$; independent variable: syntactic type, dependent variable: number of eyebrow rises).

Restrictive relative clauses are only marked with a single gestural boundary cue, in that they feature the highest number of hand beats (53% of these gestures are produced with this syntactic type), compared with their embedding sequence (however $p = .050$; independent variable: sequence segment (L, SC, R), dependent variable: number of hand beats) and with appositive relatives ($F(49, 52) = 1.9, p = .010$; independent variable: syntactic type, dependent variable: number of hand beats).

5. Discussion

5.1. Restrictive relative clauses

5.1.1. Verbal realisation

We have seen that RRCs are in majority (75%) fully governed by the main verb phrase (i.e., the main verb phrase determines their position and imposes functional restrictions, such as tense, mood, negation, and/or agreement), integrated in the nucleus. Restrictive relative clauses do not make up independent speech acts (i.e., illocutionary units) either. This configuration is illustrated in Excerpt 3:

- | | | |
|--------------|----|---|
| (3) Michelle | L | (h) and so she # disowned everything |
| | Sc | that # <i>she could associate</i> my nana # |

R for example her accent # and #

The predicate ‘could associate’ is under the scope of the main predicate in L (‘disowned’): SC and its antecedent (‘everything’) form a single argument in L’s predicative relation, fulfilling a direct object function. L and SC also make a single illocutionary unit, in that SC’s propositional content cannot be questioned independently from L (*‘no, everything that she couldn’t associate your nana’): its truth value, or relevance, is guaranteed through that of L (Debaisieux, 2016). Excerpt (3) then forms a single macro-syntactic unit.

5.1.2. Prosodic design

Restrictive relative clauses show a distinct duration pattern. This short duration is shown in Excerpt 4, in an argumentation about mistakes on an exam paper:

(4) Rhianna	L	cos hem (h) in the # test	1.9s
	Sc	they gave us	0.7s
	R	there were actually m- # grammar mistakes #	2.3s

SC is the shortest segment, referentially refining ‘the test’ at the end of L. The numerous pauses before the antecedent are evidence of Rhianna’s striving to plan the forthcoming clausal complex as a whole, in order to group the information units. The relevance of these information units would not be optimal if they happened to be split, and the small-scale discourse purpose (which consists in describing the specific subject containing mistakes) would not be fulfilled. Rhianna would also risk losing her speech turn. The items ‘test they gave us’ are grouped in a single tone-unit.

5.1.3. Visual realisation

Restrictive relative clauses are only marked with a single gestural boundary cue: they feature the highest number of hand beats. Excerpt 5 is the beginning of a description, associated with Figure 1, where (a), (b), (c), and (d) correspond to different moments in its production. Michelle has just mentioned a film Zoe has already seen, more specifically pointing out the moment she saw the film on television:

(5) Michelle	L	[(a)but i put it on the bit
	Sc	where hum (cough) (h) they (h) they were uh #] [(b) in the] [(c) garden]
	R	[(d) and they were talking]

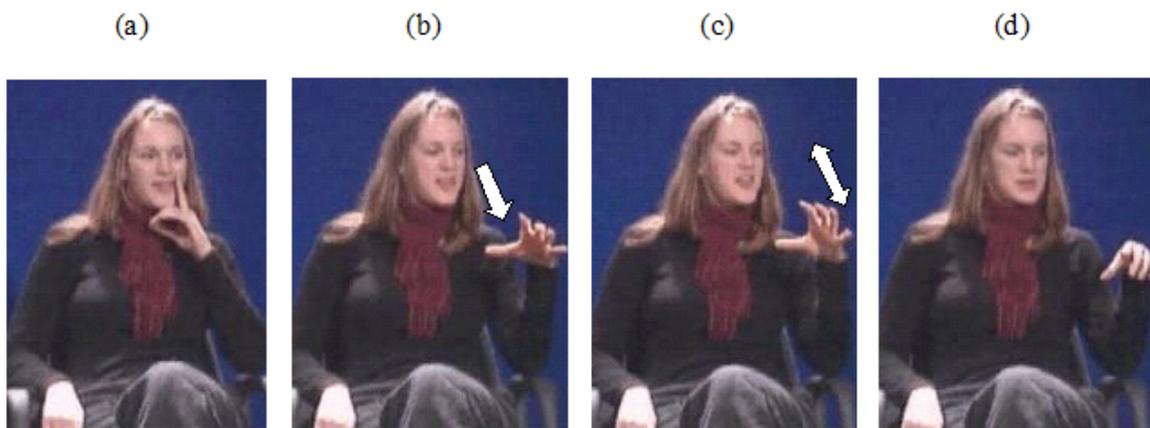


Figure 1: Two hand beats produced in correlation with a restrictive relative clause in Excerpt 5

SC stands out from the rest of the sequence through its two successive hand beats (b) and (c). The co-occurrence of these hand beats with ‘in the garden’ pragmatically indexes the most relevant informational content in the sequence (Cassell & McNeill, 1990), which is marked as the retrieval of a substantial search. Interestingly, no other beat gesture is produced in the whole discourse sequence. The palm-down open hand configuration of the hand beat takes an abstract deictic value as Michelle strives to locate an exact scene in time. Michelle partially retracts her palm-down open hand in R (d), dropping her wrist to find a new rest position for the next utterances.

5.1.4. Summary

Restrictive relative clauses mostly show a dependent syntax. While these constructions are significantly shorter, their link to their co-text is strongly marked with a direct integration to the preceding tone-unit or continuation contours. Restrictive relative clauses feature a single gestural boundary marker, as they show the greatest proportion of hand beats.

5.3. Appositive relative clauses

5.3.1. Verbal realisation

We have seen that ARCs are autonomous in majority (75%), as in Excerpt 6:

- | | | |
|---------|----|--|
| (6) Tim | L | it's only one shop for the whole thing now # |
| | Sc | <i>which is quite bad</i> # |
| | R | yeah # |

SC modifies L as a whole: its antecedent is sentential. The extract given in Excerpt 6 could be glossed without its conjunction: ‘it's only one shop for the whole thing now. It's quite bad’. SC's autonomy contradicts the descriptive reading of *which* as an introductory

marker. The co-speaker can also react to SC's content without questioning L: 'well, if you think about it, maybe it's for the best'.

5.3.2. Prosodic design

The prosodic signals indexing independence are more numerous and varied than in RRCs. We have seen that ARCs feature more extra-constituent silent pauses than the other syntactic types. Excerpt 7 illustrates this tendency:

- (7) Rhianna L even compared to Easyjet
 Sc *which is another low cost company #*
 R yeah i hate Ryanair

Rhianna sets SC apart from R with a silent pause, creating a disruption. The effect of this disruption can be seen in R, in which the realisation of a new predicate is preceded by the resumption discourse marker 'yeah'.

Appositive relative clauses are also realised with a distinct tone-unit. Likewise, they show less continuation contours than the other syntactic types. 77% of the sequences containing them do not feature any, as shown in Figure 2, representing L and SC in Excerpt 7:

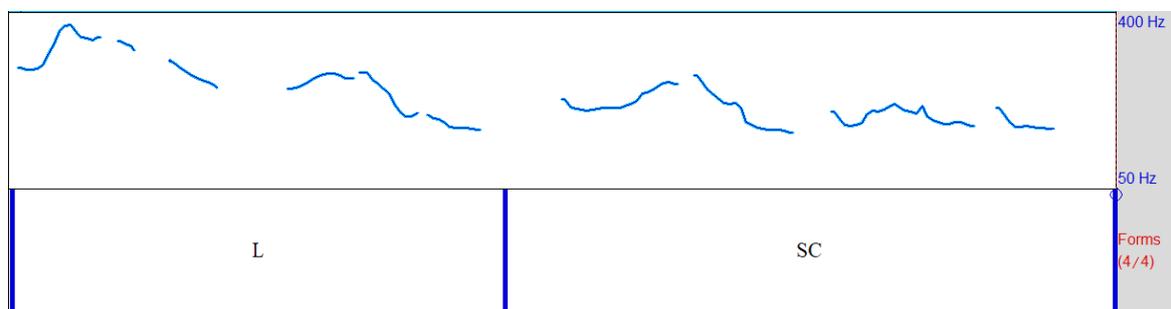


Figure 2: Intonation curve of L and SC in Excerpt 7 in Praat

While the absence of a pause between the two segments suggests their proximity at a propositional level, the definitive falling contours on L and SC indicate the two segments make two distinct discursive moves. While L supports the main theme (the staff is very rude), SC turns L into a concession, marking the common point between the two companies with a change in point of view.

5.3.3. Visual realisation

Appositive relative clauses present a distinct gesture unit, in that they feature few overlapping hand gestures between L, SC, and R. The weak distribution of overlapping hand

gestures can be seen in Figure 3 associated with Excerpt 8, where Rhianna produces a metaphorical flip with her left hand during SC, while L features a very distinct iconic gesture:

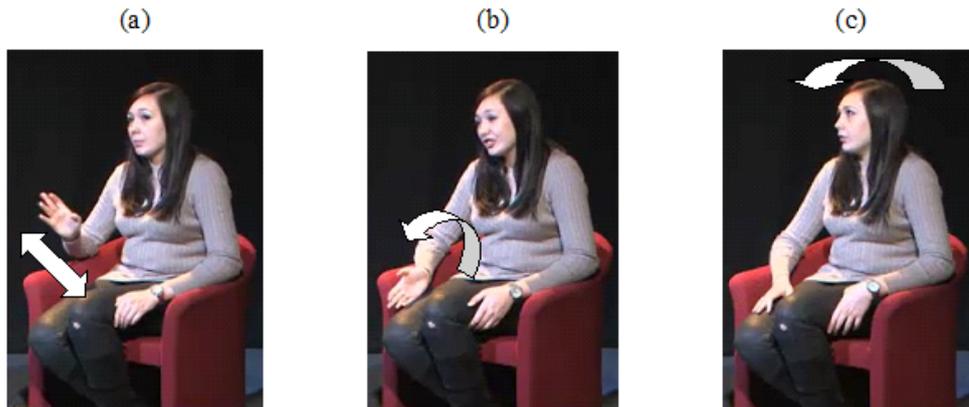


Figure 3: Metaphorical hand gesture and eyebrow rise produced in correlation with an appositive relative clause in Excerpt 8, in contrast with the iconic hand gesture produced in L and the head shake in R.

- | | |
|-------------|--|
| (8) Rhianna | i mean |
| L | [(a) my mum's pushing] me to get my license |
| Sc | (h) uh which [(b) i guess i should] # |
| R | (h) but well [(c) first of all
for the moment |

In the first image corresponding to L, Rhianna materialises her mother's advice with a sweep of her right hand corresponding to the verbal item 'pushing' (a). However, SC does not elaborate upon her mother's advice, but comments back on L's new information ('get my license'). SC introduces a change in point of view, in that the argumentation switches back to Rhianna's voice in the debate. With a head nod, Rhianna acknowledges the legitimacy of her mother's advice and marks this concession with a very distinct hand flip (b). She also raises her eyebrows in this design, taking a strong modal stance on L's arguments, and marking SC as a contrastive move. Rhianna resumes her main argumentation line in R and accompanies the next tone-units with a continuous negative head shake.

Eyebrow rises, as highlighted in Figure 3, also characteristically feature in ARCs.

5.3.4. Summary

Appositive relative clauses create a break in all three modalities considered in this study. Mainly showing total syntactic autonomy, both their rhythm and intonation create

boundaries, with silent pauses and terminal contours. Appositive relative clauses are also significantly produced with distinct hand gestures and numerous eyebrow rises.

6. Conclusion

Our analysis confirms that the two syntactic types can be distinguished in their degree of autonomy, not only from a verbal perspective but also from vocal and gestural points of view.

Restrictive relative clauses are more integrated in their (left) co-text. While showing a total verbal linkage featuring no main macro-syntactic disjunctive cue, their visual integration is also almost complete. However, restrictive relative clauses rhythmically stand out, in regard to their duration and to their production of hand beats. In this type of SC, both intonation and gesture signal dependence.

Appositive relative clauses are more independent forms, mainly showing vocal boundaries and very distinct hand gestures. Disruption is then more perceptible, with the highest distribution of extra-constituent pauses, and the weakest distribution of continuation contours. With this SC, both prosody and gesture index autonomy.

The differences regarding the distribution of the factors in the two syntactic constructions suggest that no common boundary cue is systematically used during subordination. However, the significant presence of hand beats and eyebrow rises hint at the prevalent use of prosodic gestures in both types of subordinate constructions. Interestingly, in the vocal modality, rhythmic cues also play a seminal role in the demarcation of both constructions (Lelandais & Ferré, 2016).

When analysing how semiotic units form larger sequences of action in discourse and conversation, spontaneous speech presents both complex chains of structures embedded in one another, and disruptions in which the discourse parts no longer follow one another. Subordinate constructions introduce a break when they establish a different assertive position from the preceding utterance as is the case of ARCs. While this break can directly be expressed with syntactic or discursive means, prosody creates a break immediately afterwards through rhythmic features or pitch upsteps, signalling that the previous elements have to be recontextualised. To avoid a gap between the co-speaker's representations and the speaker's input, gestures give pragmatic instructions about the informational value of the propositional content (e.g., hand beats, eyebrow rises), showing a different relation to the linearity of discourse through the use of physical space.

This study aimed to demonstrate that a composite, contextual vision of linguistic communication sheds new light on subordination, deriving from numerous interactions between verbal, vocal, and visual components. The analysis focuses on discursive demarcation; one way to further a qualified picture of subordination would be to apply the analysis to focalisation, i.e. on a thematic organisational mode between foreground and background information.

Competing interests

The authors have no competing interests to declare.

References

- Arnold, D., & Borsley, R. D. (2008). Non-restrictive relative clauses, ellipsis and anaphora. In S. Müller (Ed.), *Proceedings of the HPSG08 Conference* (pp. 325–345). Stanford, CA: CSLI Publications.
- Beattie, G. W. (1978). Sequential temporal patterns of speech and gaze in dialogue. *Semiotica*, 23(1–2), 29–52. doi:10.1515/semi.1978.23.1-2.29
- Beckman, M. E., Hirschberg, J., & Shattuck-Hufnagel, S. (2005). The original ToBI system and the evolution of the ToBI framework. In S-A. Jun (Ed.), *Prosodic Typology* (pp. 9–54). Oxford: Oxford University Press.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. London: Pearson Education.
- Bigi, B. (2012). SPPAS: A tool for the phonetic segmentations of speech. In N. Calzolari (Ed.), *Proceedings of the International Conference on Language Resources and Evaluation* (pp. 1748–1755). Luxemburg: European Language Resources Association.
- Boersma, P., & Weenink, D. (2013). Praat: Doing Phonetics by Computer [Computer Software]. Retrieved from <http://www.fon.hum.uva.nl/praat/>
- Bolinger, D. (1984). Intonational signals of subordination. In C. Brugman & M. Macaulay (Eds.), *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* (pp. 401–413). Berkeley, CA: eLanguage.
- Borsley, R. D. (1992). More on the difference between English restrictive and non-restrictive relative clauses. *Journal of Linguistics*, 28(1), 139–48. doi:10.1017/S0022226700015024
- Bressemer, J., & Ladewig, S. H. (2011). Rethinking gesture phases: articulatory features of gestural movement? *Semiotica*, 184, 53–91. doi:10.1515/semi.2011.022
- Burton-Roberts, N. (1999). Language, linear precedence and parentheticals. In P. Collins & D. Lee (Eds.), *The Clause in English* (pp. 33–51). Amsterdam: John Benjamins.
- Calbris, G. (2011). *Elements of Meaning in Gesture*. Amsterdam: John Benjamins.

- Cassell, J., & McNeill, D. (1990). Gesture and ground. In K. Hall, J-P. Koenig, M. Meacham, S. Reinman, & L. A. Sutton (Eds.), *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* (pp. 57–68). Berkeley, CA: eLanguage.
- Cavé, C., Guaitella, I., Bertrand, R., Santi, S., Harlay, F., & Espesser, R. (1996). About the relationship between eyebrow movements and F0 variations. In H. T. Bunnell & W. Idsardi (Eds.), *Fourth International Conference on Spoken Language* (pp. 2175–2178). Philadelphia, PA: Institute of Electrical and Electronics Engineers.
- Chafe, W. (1988). Linking Intonation Units in Spoken English. In J. Haiman & S. A. Thompson (Eds.), *Clause Combining in Grammar and Discourse* (pp. 1–27). Amsterdam: John Benjamins.
- Cotte, P. (2008). Les propositions relatives et l'énonciation. *Cycnos*, 17. Retrieved from <http://revel.unice.fr/cycnos/index.html?id=1697>
- Couper-Kuhlen, E. (1986). Intonation and grammar. In E. Couper-Kuhlen (Ed.), *An Introduction to English Prosody* (pp. 139–157). Tübingen: Max Niemeyer Verlag.
- Crystal, D. (1969). *Prosodic Systems and Intonation in English*. Cambridge: Cambridge University Press.
- Debaisieux, J-M. (2016). Toward a global approach to discourse uses of conjunctions in spoken French. *Language Sciences*, 58, 79–94. doi:10.1016/j.langsci.2016.04.001
- De Kok, I., & Heylen, D. (2009). Multimodal end-of-turn prediction in multi-party meetings. In J. L. Crowley, Y. Ivanov, & C. Wren. *Proceedings of the 2009 International Conference on Multimodal Interfaces* (pp. 91–98). New York: Association for Computing Machinery.
- De Vries, M. (2006). The syntax of appositive relativization. *Linguistic Inquiry*, 37(2), 229–270. doi:10.1162/ling.2006.37.2.229
- Emonds, J. E. (1979). Appositive relatives have no properties. *Linguistic Inquiry*, 10, 211–243.
- Enfield, N. (2009). *The Anatomy of Meaning: Speech, Gesture and Composite Utterances*. Cambridge: Cambridge University Press.
- Fabb, N. (1990). The difference between english restrictive and nonrestrictive relative clauses. *Journal of Linguistics*, 26(1), 57–77. doi:10.1017/S0022226700014420
- Fox, B. A. & Thompson, S. A. (2009). On formulating reference: an interactional approach to relative clauses in english conversation. *Papers in Pragmatics*, 4(1), 183–196.

- Gilbert, A., Itzhak, I., & Baum, S. (2016). A cross-language investigation of word segmentation by bilinguals with varying degrees of proficiency: Preliminary results. In J. Barnes, A. Brugos, S. Shattuck-Hufnagel, & N. Veilleux (Eds.), *Proceedings of Speech Prosody 2016* (pp. 36–39). Boston, MA: International Speech Communication Association.
- Granström, B., & House, D. (2005). Audiovisual representation of prosody in expressive speech communication. *Speech Communication*, 46(3), 473–484.
doi:10.1016/j.specom.2005.02.017
- Goodwin, C. (1979). The interactive construction of a sentence in natural conversation. In G. Psathas (Ed.), *Everyday Language: Studies in Ethnomethodology* (pp. 97–121). New York: Irvington.
- Haegeman, L. (2010). The internal syntax of adverbial clauses. *Lingua*, 120, 628–648.
doi:10.1016/j.lingua.2008.07.007
- Haiman, J., & Thompson, S. A. (1984). ‘Subordination’ in Universal Grammar. In C. Brugman & M. Macaulay (Eds.), *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* (pp. 510–523). Berkeley, CA: eLanguage.
- Halliday, M. A. K. (1985). *An Introduction to Functional Grammar*. London: Edward Arnold.
- Halliday, M. A. K. & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Heringa, H. (2007). Appositional constructions: coordination and predication. In M. Kluck & E-J, Smits (Eds.), *Proceedings of the Fifth Semantics in the Netherlands Day* (pp. 67–82). Groningen: University of Groningen Press.
- Hirschberg, J. & Grosz, B. (1992). Intonational features of local and global discourse structure. In M. P. Marcus (Ed.), *Proceedings of the Workshop on Speech and Natural Language* (pp. 441–446). Morristown, NJ: Association for Computational Linguistics.
- Hirst, D. (2007). A Praat plugin for Momel and INTSINT with improved algorithms for modelling and coding intonation. In J. Trouvain & W. J. Barry (Eds.), *Proceedings of the XVIth International Conference of Phonetic Sciences*. Retrieved from <http://www.icphs2007.de/>
- Holler, A. (2005). Expressing communicative-weight assignment discourse structurally. In C. Sassen, A. Benz, & P. Kühnlein (Eds.), *Proceedings of the Workshop of Constraints on Discourse* (pp. 88–94). Dortmund: Universität Dortmund.

- Holler, J., Schubotz, L., Kelly, S., Hagoort, P., Schuetze, M., & Özyürek, A. (2014). Social eye gaze modulates processing of speech and co-speech gesture. *Cognition*, 133(3), 692–697. doi:10.1016/j.cognition.2014.08.008
- Huddleston, R., & Pullum, G. K. (2002). *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Jackendoff, R. (1977). *X-Bar Syntax: A Study of Phrase Structure*. Cambridge, MA: MIT Press.
- Jespersen, O. (1927). *A Modern English Grammar on Historical Principles*. London: George Allen & Unwin.
- Kendon, A. (1972). Some relationships between body motion and speech. In A. W. Siegman & B. Pope (Eds.), *Studies in Dyadic Communication* (pp. 177–210). New York: Pergamon.
- Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.
- Kipp, M., Neff, M., & Albrecht, I. (2007). An annotation scheme for conversational gestures: how to economically capture timing and form. *Language Resources and Evaluation*, 41, 325–339. doi:10.1007/s10579-007-9053-5
- Kita, S., van Gijn, I., & van der Hulst, H. (1998). Movement phases in signs and co-speech gestures, and their transcription by human coders. In I. Wachsmuth & M. Frölich (Eds.), *Gesture and Sign Language in Human-Computer Interaction Proceedings* (pp. 23–35). Berlin: Springer.
- Krifka, M. (2007). Basic notions of information structure. *Interdisciplinary Studies on Information Structure*, 6, 13–55.
- Lambrecht, K. (1996). *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*. New York: Cambridge University Press.
- Langacker, R. W. (1991). Cognitive grammar. In F. G. Droste & J. E. Joseph (Eds.), *Linguistic Theory and Grammatical Description: Nine Current Approaches* (pp. 275–306). Amsterdam: John Benjamins.
- Langacker, R. W. (2008). *Cognitive Grammar. A Basic Introduction*. Oxford: Oxford University Press.
- Lascarides, A., & Stone, M. (2009). Discourse coherence and gesture interpretation. *Gesture*, 9(2), 147–180. doi:10.1075/gest.9.2.01las

- Lehmann, C. (1988). Towards a typology of clause linkage. In J. Haiman & S. A. Thompson (Eds.), *Clause Combining in Grammar and Discourse* (pp. 181–225). Amsterdam: John Benjamins.
- Lelandais, M., & Ferré, G. (2016). Prosodic boundaries in subordinate syntactic constructions. In J. Barnes, A. Brugos, S. Shattuck-Hufnagel, & N. Veilleux (Eds.), *Proceedings of Speech Prosody 2016* (pp. 183–187). Boston, MA: International Speech Communication Association.
- Local, J. (1992). Continuing and restarting. In P. Auer & A. Di Luzio (Eds.), *The Contextualization of Language* (pp. 273–296). Amsterdam: John Benjamins.
- Local, J. (2007). Phonetic detail and the organisation of talk-in-interaction. In J. Trouvain & W. J. Barry (Eds.), *Proceedings of the XVIth International Conference of Phonetic Sciences*. Retrieved from <http://www.icphs2007.de/>
- Longacre, R. E. (1985). Sentences as combinations of clauses. In T. Shopen (Ed.), *Language Typology and Syntactic Description: Complex constructions* (pp. 372–420). Cambridge: Cambridge University Press.
- Lytvynova, M. & Dao, H. L. (2014). Les relatives appositives entre intégration syntaxique et intégration discursive. In *4^e Congrès Mondial de Linguistique Française*. Retrieved from <https://halshs.archives-ouvertes.fr/halshs-01506751/>
- Matthiessen, C., & Thompson, S. A. (1988). The structure of discourse and ‘subordination’. In *Clause Combining in Grammar and Discourse* (pp. 275–329). Amsterdam: John Benjamins.
- McNeill, D., & Levy, E. T. (1993). Cohesion and gesture. *Discourse Processes*, 16, 363–386. doi:10.1080/01638539309544845
- McNeill, D., Levy, E. T., & Duncan, S. D. (2015). Gesture in discourse. In D. Tannen, H. E. Hamilton & D. Schiffrin (Eds.), *The Handbook of Discourse Analysis* (pp. 262–289). Oxford: John Wiley & Sons.
- McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. Chicago, IL: University of Chicago Press.
- McNeill, D. (2005). *Gesture and Thought*. Chicago, IL: University of Chicago Press.
- Melis, G. (2008). Relatives et types de qualification. *Cycnos*, 17. Retrieved from <http://revel.unice.fr/cycnos/index.html?id=1693>

- Muller, C. (2008). La relation au verbe principal dans les relatives prédicatives en français. *Faits de Langue*, 31–32, 337–346.
- Norris, S. (2004). *Analyzing Multimodal Interaction: A Methodological Framework*. New York: Routledge.
- Peterson, P. (1999). On the boundaries of syntax: Non-syntagmatic relations. In P. Collins & D. Lee (Eds.), *The Clause in English* (pp. 229–250). Amsterdam: John Benjamins.
- Potts, C. (2005). *The Logic of Conventional Implicatures*. New York: Oxford University Press.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. London: Longman.
- Rossano, F. (2012). *Gaze Behavior in Face-to-face Interaction* (Doctoral dissertation). Nijmegen: Max-Planck Institut for Psycholinguistics. Retrieved from <http://hdl.handle.net/2066/99151>
- Sloetjes, H., & Wittenburg, P. (2008). Annotation by category: ELAN and ISO DCR. In N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis, D. Tapias. (Eds.), *Proceedings of the 6th International Conference on Language Resources and Evaluation*. Retrieved from <http://www.lat-mpi.eu/tools/elan/>
- Smessaert, H., Cornillie, B., Divjak, D., & van den Eynde, K. (2005). Degrees of clause integration: from endotactic to exotactic subordination in dutch. *Linguistics*, 43(3), 471–529. doi:10.1515/ling.2005.43.3.471
- Streeck, J. (2009). *Gesturecraft. The Manufacture of Meaning*. Amsterdam: John Benjamins.
- Sweetser, E. (2006). Looking at space to study mental spaces: co-speech gesture as a crucial data source in cognitive linguistics. In M. Gonzalez-Marquez, I. Mittleberg, S. Coulson & M. Spivey (Eds.), *Methods in Cognitive Linguistics* (pp. 203–226). Amsterdam: John Benjamins.
- Thompson, S. A. (2002). “Object complements” and conversation: Towards a realistic account. *Studies in Language*, 26(1), 125–163. doi:10.1075/sl.26.1.05tho
- Tomlin, R. S. (1985). Foreground-background information and the syntax of subordination. *Text*, 5(1–2), 85–122. doi:10.1515/text.1.1985.5.1-2.85

- van Rijn, M. A. (2017). *The Expression of Modifiers and Arguments in the Noun Phrase and Beyond* (Doctoral dissertation). Amsterdam: University of Amsterdam. Retrieved from <https://dare.uva.nl/search?identifier=2228b1a1-835e-4fde-9795-dd2df3e49a5a>
- van Valin, R. D. (1984). A typology of syntactic relations in clause linkage. In C. Brugman & M. Macaulay (Eds.), *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* (pp. 542–58). Berkeley, CA: eLanguage.
- Wells, J. C. (2006). *English Intonation: An Introduction*. Cambridge: Cambridge University Press.
- Wennerstrom, A. K. (2001). *The Music of Everyday Speech. Prosody and Discourse Analysis*. Oxford: Oxford University Press.