Replication as speech enactment marker

Auteur : Boland, Raphaël
Promoteur(s) : Perrez, Julien
Faculté : Faculté de Philosophie et Lettres
Diplôme : Master en linguistique, à finalité approfondie
Année académique : 2019-2020
URI/URL : http://hdl.handle.net/2268.2/9450

Raphaël Boland	2019-2020
Travail de fin d'études	Promoteur: Julien Perrez

Replication as speech enactment marker

Table des matières

Introduction

Repetition occurs at all levels of language, from basic morphemes to entire stretches of discourse. Everyday conversations are interspersed with reproductions or variations of what has been previously said, whether they echo our own words or the words of others. This ubiquitous phenomenon proves to be essential in the flow of any interaction. It is not only used between adult speakers, but is also massively found in language development (Bennett-Kastor 1994), in second language acquisition (Tomlin 1994), or in sign language (Bergman and Dahl 1994). Whenever the communication is limited by restricted linguistic resources, repetition is used as a pragmatic lever that counterbalances the lack of shared code. In this thesis, three kinds of repetition will come under scrutiny:

- **Reported speech (RS)/ Constructed speech (CS):** RS is a common tool in interaction. Through direct or indirect speech, we report fragments of previous utterances. Those fragments can either be creations made to bring our speech into life, or verisimilar reproductions. When the RS is blatantly devised at the very moment of the speech event (for some pragmatic purposes), I will label it "constructed speech"; but when it is supposed to report more or less faithfully something that has been said in a previous conversation, the term "reported speech" remains appropriate. Typically (though it is far from being always the case), CS are introduced with the so-called "new quotative" (Buchstaller 2001) *like* (I was like: "…"). Note that RS and CS are two opposite poles on a continuum and refer therefore to a same phenomenon. The distinction between the two is purely methodological and does not categorize them in two separate conceptual fields.

  Aside from (in)direct RS, a third possibility is to *enact* (not only *report*) the speech, i.e. to animate words, for instance during a joke session. In the course of this work, the notion of "enacted reported speech" will prove to be more convenient for the subject matter.

- **Repetition (RE):** Within the scope of this study, RE will refer to the doubling of a word in conversation, e.g. *a blue, blue sky* or *These days are long long gone…*, or the immediate repetition of someone's previous utterance, e.g. *This painting is ugly – ugly?* By repeating, a speaker or listener tries to convey his attitude towards the proposition. This piece of information will then be received by the addressee, who, on basis of the pragmatic setting, the supposed communicative intentions of the sender, and other resources (e.g. cultural knowledge), will try to interpret it the best way he can.

- **Reduplication (RD):** reduplicated units pertain to grammaticalized/lexicalized RE (Dansieh 2011), e.g. Jamaican Creole *redi-redi*, "red-spotted" (Fischer 2011), Japanese *samazama*, "various" (Kauffmann 2015), Urdu *garam-garam*, "nicely hot" (Montaut 2009). I will divide them into two branches: "diagrammatic" RD, which modulates quantitatively and qualitatively the core notion of the simplex (i.e. the initial meaning of the simplex is enriched with a specific aspect: intensification, vagueness, plurality, specificity, distribution, …), and "imagic" RD, centered on phonetic musicality (i.a. rhyming RD and ablaut RD).

The issue I will try to address revolves around the contextual regulation and negotiation of expressive meaning through the use of replication (RP; a collective label for RS, RE, and RD, after the example of Dingemanse 2015). The notion of "expressiveness" is borrowed from Jakobson's functions of language, where the pole of the sender is called "expressive function" (Jakobson 1960). Expressiveness is an umbrella term designating everything that relates a speech act to the attitudes, intentions and feelings of a speaker. Underneath the words, there is always a language user who wants to express his communicative intention *to* someone else (if there was not a recipient, utterances would not serve any purpose). A speech act includes a twofold dimension of meaning: the encoded, informative meaning, and the expressive meaning, which pertains to the communicative intention of the speaker with regard to the objective he is willing to reach, i.e. the purpose of his act.

I hypothesize that RP is by nature a device foregrounding a speaker's expressive meaning. For instance: if, during a discussion, interlocutor A makes a wrong interpretation of interlocutor B's words, B will immediately react by saying *No, no [I don't mean that]*, and not simply *no [you're wrong]*. The most straightforward explanation for that would consist in the claim that repeating rests on the principles TWO IS MORE and MORE IS INTENSIVE. It is nevertheless questionable that B, while saying *no, no*, insists on the fact that A is wrong, viz. not **[I didn't mean that]*, but *[you're totally wrong]*. He rather aims at defeating a wrong implicature inferred by A, and, by repeating, makes clear that he has something else in mind. The communicative intention (= expressive meaning) is foregrounded and "manually" regulated ("You shouldn't understand what I say like *this*, but like *that*"). Normally, A should understand after B's RE that he actually has to infer a second, at first less preferred option.

If these two equations are brushed away, what is left to explain the connection between RE (and, more broadly, RP) and regulation of expressive meaning? After introducing RS and quotatives (point 2.1 and 2.2), I will compare RS and RE and show that they are both pragmatic devices of the same nature (point 2.3), both marking the fact that the speech is enacted (notion explained in

chapter 1). First, it will be observed that both of them orient inferences. A participant can make clear that he intends to communicate something through defeating unwanted potential inferences, thereby directing the partner to the right "expressive" meaning (cf. *No no*). On a further step, it will be observed that far from all RE are based on inference constraint. To set it clear, I will propose a four-part division of RE underlying the general concept of *echo* (chapter 3). An *echo* means that, with the help of RE, speech participants find a common interface through which they can interpret each other's expressive meaning. Instead of manipulating inferences, i.e. constantly directing each other to the right interpretive groove (it is potentially exhausting to specify all the time: "You shouldn't understand what I say like this, but like that"), they organize themselves in order to be on the same wavelength, so that nobody moves away from the shared conversational trajectory.

In the following chapter (chapter 4.), I will try to disclose a potential link between diagrammatic RD and echoes, and what ultimately distinguishes them. Consequently, I will make an attempt to connect echoes (which occur spontaneously in dialogue) and entrenched reduplicated forms, without any cleaving presuppositions about what is strictly lexical/grammatical and what is strictly conversational. Indeed, I argue that RP is above all a pragmatic tool which originally does not hold any conceptual value. For example: although diagrammatic RD goes most of the time together with a nuance of addition, duration, or intensification, I will not support the idea that such "increase"-type of meaning is an inherent semantic feature of RD, but that it results from defining traits of RP, especially its relation to the negotiation of expressive meaning in context. In diagrammatic RD, this relation manifests itself in the form of a perceptual alignment that I will call *sharing of presence*. Modulation of the core notion (i.a. the above-mentioned "increase") is directly connected to this way for two partners to access a shared perceptual viewpoint. In fact, instances of RD, as much as RE, cannot be analyzed on their own and hinge more than anything on contextualized utterances and attitudes of speakers towards their activities or topics of discussion.

This thesis is supposed to be a non-experimental, deductive work, based on previous research on the topic and observations of everyday conversations. Material of discussion will be fragments of conversations, mostly in French and English, and, for demonstrative purposes only, from a variety of other languages (as Japanese or Finnish). The cross-linguistic orientation given to this work is meant to underscore the fact that what will be discussed is not language-specific, but generic. To my knowledge, RS, RE, and RD are a constitutive part of every language, and what is more, they show a noteworthy functional and semantic consistency throughout the world. By integrating elements of other languages, I intend to show that my concern goes beyond French and English (which will be considered as models pointing to a higher degree of generality), and encompasses every instance of RP, regardless of the language in which they occur.

Fragments in French are systematically drawn from the CLAP-database (Corpus de LAngues Parlées en interaction). English conversations and examples come for the most part from academic papers. There will be also many coined examples, whose function is to assist the reasoning and make a bit more concrete the (sometimes abstract) ideas that will be proposed. Coined examples are not considered as "data": their role is purely argumentative and/or illustrative.

Before moving on to the heart of the matter, I will first introduce the conceptual frame that will ground this research about RP, the so-called "distributed" view on cognition, which has the capacity to explain how expressive meaning is negotiated in interaction. The core notion in this work, *speech enactment*, will derive from this particular cognitive stance, what makes a sub-chapter necessary before beginning.

1. Speech enactment as contextual negotiation of expressive meaning

What is the purpose of RP? How does it reconfigure the speech so that its expressive meaning becomes highlighted and negotiated in a specific context? An answer could be that RP enables the speaker to enact his speech. "Enacting" is a term which is very often found in phenomenological and embodied research on language, more specifically in distributed cognition (see for instance Di Paolo & al. 2018; Bottineau 2013). Distributed cognition takes for certain that cognition is not "inside the head", but evolves through our interrelationship with the environment. Cognition is thus distributed into our environment and defined through the way we move in it, and consequently does not boil down to an array of neuronal configurations. The particularity of humans is that their ecological space is semiotized. It means that cognition is not (or not exclusively) forged through our immediate interaction with our surroundings, but is mediated by a history of coordinated dialogic acts with other agents, subsequently regulated by culturalized patterns. In this regard, language cannot be a window on cognition, i.e. it cannot be transcendentally explained through processes operating "upstream" in the mind, but *is* cognition, because it reflects the way we interact and interpret each other through norms developed or acquired on a local and cultural level (for a full account on distributed-enacted cognition, see Penelaud 2010).

In this context, "enactments" designate the permanent actions an agent takes to relate to the world he lives in or to adapt his relation with it. In this thesis, I will take inspiration from that view on cognition. Accordingly, "speech" enactments will refer to linguistics devices that affect the dynamics of the dialogue by changing ("adapting") the mode of interpretation, e.g. when a common conversation develops into a joke session or when participants begin to speak figuratively (irony).

While joking or making irony, one namely displays one's expressive meaning (because if there is no hint at the fact that a comment is intended to be a joke and the interlocutor takes it seriously, this is not a joke any more). In the same way, any RE like *no no [I don't mean that]* thematizes the speaker's intention: on a first degree, *no no* is only a denial, but on a second one, it brings to the surface the way an utterance is to be interpreted. Adding a "second degree" (= foregrounding the intention/expression) is an explicit indication on how an interactive move (not necessarily linguistic) is to be understood by the partner. If foregrounding an expressive meaning through an enactment is a way to steer the interpretation, and consequently, to regulate a dialogue, then enactments are synchronizing devices. In other words, they involve all participants by making them share a common interpretative stance about what is spoken about. It is thus a useful tool to resolve misinterpretation or to ensure that everybody share a common point of departure (i.e. a thematical anchor point). Most of the time, expectations inferred from previous encounters (habits) are enough to guess someone else's attitude or intentions behind a specific utterance, but sometimes it fails and collaboration is necessary for a conversation not to grind to a halt.

One could confuse *speech enactment* with *speech act*. Indeed, one could say that every speech act is expressive (it communicates the speaker's attitude), but this notion is much too broad and differs from the phenomenon I want to cover. First of all, a speech act includes what one does by saying something: answering, promising, asking, apologizing, complaining, refusing, congratulating, etc. A speech enactment, on the other hand, relates to everything that could serve as an answer to the question *What do you mean?*, e.g. *A. We have to go now/ B. What do you mean: now-instantly or now-in five minutes?/ A. Now instantly, of course*, or *A. I'm very fond of you/ B. What do you mean?/ A. Well… that I'm "very" very fond of you[1]*. In these examples, the enactment is the answer to the *what do you mean?* On the contrary to a speech act, a speech enactment systematically draws attention to the meaning expressed. Another example with a RS, similar to the RE of *No* presented in the introduction (that I take here in the very broad sense of "framing the speech" by a performative verb pointing out to the illocution. Explications for this choice are found in the next chapter): *[A teacher speaking to his pupil] I warn you: if I see you again with a mobile phone in the hands during class, I take it and throw it through the window*. The thematized aspect is not the illocutionary act (a warning), but the expressive meaning: *"you think maybe that in the XXIth century, a teacher cannot deteriorate the belongings of his students any more if he thinks they deserve it [wrong inference], but believe me: I will certainly do it [right inference]"*. This last example offers a second distinction: a speech enactment is always embedded in a living, embodied communicative context, as well as a sociocultural one (a culture in which teachers cannot

---

1   When no source is given, the example is coined.

deteriorate the belongings of their pupils anymore). Outside a context, without taking into account that the teacher speaks to a pupil who may think that he might not realize his warning on the basis that such thing does not happen anymore in the 21st century, a speech enactment is not understandable. In isolation, the sentence *I warn you: if I see you…* is opaque (the "I warn you" cannot be interpreted). A speech act, on the contrary, can be understood outside a fully embodied context: if I read on a post-it stuck on a chair "Don't touch it.", I know it's an order requesting that nobody touches the chair; the illocutionary act can be understood without a full linguistic engagement. But if the post-it is stuck on a chair and a cupboard at the same time, I will wonder what the author meant by that, and after a moment of thought, say to myself: "He certainly meant the chair, because it looks more fragile". While being the recipient, I also jump into the addresser's mind (the *he*). A full situation of communication is created, because I have to visualize him and to guess his intentions, as if he was present. If there was not any ambiguity, the speech act would be perfectly understandable and no real situation of communication would be created. Only the message (the statement), not the person who wrote the post-it (the enunciator), would have been the focus of attention. A speech enactment, on the other hand, demands a co-presence between participants, even if one of them is absent.

This example leads us to a third distinction: by successively playing both roles at the same time, one propels the whole communication situation from the here-and-now of the conversation to a virtual stage situated beyond the here-and-now (an "enacted" stage). Or, to put it simply, one shifts from point of view and understand things from a determined perspective. Imagination must be put to use to set up a whole context. In the post-it example, imagining the author of the message requires taking perspective on oneself in order to understand the intention of somebody else. Maynard (2002) called that "perspective of becoming" (in the quote, "reader" should be replaced by "receiver" in the broadest sense):

> To understand how a person feels […], one needs to guess at the person's intention, feelings, and emotion, and to experience it oneself […] Taking this <perspective of becoming> requires that the reader attempts to generate the psychological state of someone else, as if the reader were that person. (p.93)

The "virtual stage" that is reached through the perspectivization process is plainly suggested by the word "enactment". One could also use the term "animation" ("speech animation"), but its denotation is too restricted and does not accurately describe some aspects of speech enactments (in particular echoes and diagrammatic RD), although it accounts well for some others (irony, joking). It is also worth noticing that, even if situated on a stage (i.e. even if things are attended from a different perspective), the reference point is always the same, enactment or not, what entails that the

deictical space remains unchanged (e.g. pointing to something that is not there, but imagined to be there).

In the post-it example, both (present or absent) participants are still differentiated, but they are also situations where the producer interprets not somebody else, but himself. This happens most of the time when so-called "fillers" are used: *I mean, like, well, etc.* While speaking, one has to constantly check if successive words, ideas, topics are cohesive and coherent, in other words, that they can be integrated syntactically and practically in the ongoing speech, as if on a loop:

Interpretation Int.

Prod.uction Prod.

…
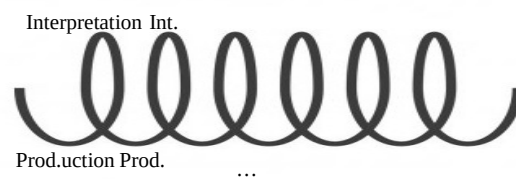
A concrete example of reflexive enacted speech would be a teacher who, during a lesson, faces the confused look of some of his pupils and understands that he should redraft his ideas. Therefore, he raises the eyes for a few seconds, and, with urging gestures (supposed to boost his stream of thought), says "let's say, well, I mean…" and then goes on. During this short time, he cuts himself off and adopts a split attitude to rearrange his thoughts. As long as he spells out what he "means", the speech is enacted. In these situations, it is not a *I* who speaks, it is a *me* (i.e. a perspectivized *I*). In a distributed perspective, there is no *I*, because the sense of self is distributed in the underlying network, but rather a *me* open to interactions, interfacing between the world and the self (Varela 1996). The "perspectivized *I*" could be related to the notion of *body image*[2]. In psychological terms, the latter refers to the self-awareness of one's own body. A complementary notion of *body image* is *body scheme.* A scheme means that in our daily life, we act in the world according to habitual, integrated clusters of unconscious, habitual moves (riding a bike, swimming, sneezing in a handkerchief, grasping a glass, etc.). When we are engaged in non-habitual actions, such as learning a piece of music or combining a string of utterances, body images are activated. When learning a piece, namely, a musician has to check permanently if his fingers are in the right place in the right moment and becomes consequently aware of his hands (Kim 2020). A lecturer, through his gestures or the regular use of fillers, most likely also creates an image of himself to interact with. Body images participate in enactments, because they help speakers to self-interpret (and not only interpret

---

2   See for example Gallagher (1986) for conceptual clarification.

in the absence of the interlocutor, as in the post-it example above) their speech act when they face, for instance, misinterpretation, or to guarantee self-control over their production.

In mental disorders such as schizophrenia, the body image processing is impaired: the ailing does not feel his physical contours, and cannot therefore "feel" himself speaking through body images. This leads to what Kuperberg labels a "derailment" (Kuperberg 2010) leading to a "loosening of associations". Coherence and cohesiveness are disturbed, because the schizophrenic cannot interpret his speech ("in very severe cases, positive thought disorder manifests as unintelligible speech in which neither the individual words nor the sentences being strung together seem to correspond to any discernable overall meaning", p.4). So, generating coherent talk or thought needs self-representations (body images) propelling the whole body in a second speech situation situated in a there and then where self-awareness is enabled.

Lastly, the content osf a speech enactment is *(re)presentational,* while the content of a speech act is *propositional*. *Propositional* (one could also say *digital*) entails that the meaning of the contextualized utterance is symbolic, conventionalized, descriptive, and has only an exchange value (the meaning "interdiction" is simply coded and decoded in the imperative "Don't do that."). (*Analogic re-)presentation,* on the other hand, entails a live performance: some instances of speech enactment, like personal echoes or RS introduced by personal quotatives in the present tense (a performative such as: I say: "...", you say "..."), present an utterance, e.g. the utterance "I tell you: don't do that" presents an *I* saying: "don't do that" in a framed context corresponding exactly to the speech event. On the contrary, a RS framed by a quotative in the preterite (*He said: "…"*), *re*presents a past situation. Most onomatopoeia, like *blah blah*, *woof-woof* or *wham!,* also represent or mimicry voices situated outside the conversation.

Preference to enact a speech or to use regular speech acts with a minimum of mutual interpretive effort depends on the ability of the speaker to mold his speech into prefabricated discourses, which function as sources of stability in our daily interaction. Peirce called such cultural habits *final* or *normal interpretants* (Savan 1980). Final interpretants condition interpretation so as to avoid permanent negotiation about what is the meaning of every produced sign. For instance, a baby or a little child, who has not been sufficiently immersed in a community of meaningful practices (and has consequently not yet developed enough "normal interpretants"), cannot interpret his words on his own. The interpretation is built in intersubjectivity and in a playful manner with i.a. the use of RE or echoes (Bennett-Kastor 1993). By contrast, an adult speaker is usually aware that his speech is conform to prefabricated discourses (cultural habits). In that case, there is no need to enact.

I've drawn below a (provisional) table listing the standards of comparison between speech acts and speech enactments:

Speech act	Speech enactment
Here and now	There and then (perspectivization)
Immediate interpretation	Emerging interpretation
Expressive meaning not foregrounded	Expressive meaning foregrounded
Propositional/digital	(Re)presentational/analogic
Language as product	Language as interactional achievement

Actually, a "speech" enactment is much more an embodied phenomenon than a linguistic one: every interaction displays a rich dynamics of quick coordinating movements testifying the continuous effort to give meaning to the second-by-second unfurling dialogic act. Turning up the microscope, synchronizing pico-scalar vocal and bodily dynamics (postural orientation, shift of gaze, falling or rising pitch, sound lengthening or shortening, etc.) can be uncovered that reveal the constant meaning negotiation (or perspective sharing) between two agents (Thibault 2011). Although most of the time pre-reflexive, such coordinating acts are always anchored in our sensorimotor apparatus and based on the historical and cultural experience of each individual.

So, what is actually *not* enacted? Typically, lexicogrammatical patterns and entrenched schematic constructions, unless, once again, interpretation issues come up (A: Are you all right? B: *Silence*. A: What did I say wrong?). Every interaction is precarious, even the most natural ones (such as greeting), and thus likely to be enacted. After the example of Thibault (2011), I will label stabilized linguistic patterns "second-order languaging", and the real-time, embodied, dynamical negotiation of meaning "first-order languaging". Researchers in distributed cognition usually use "languaging" instead of "language", because the latter suggests that each speaker has stable linguistic skills that he extracts at the right moment (Bottineau 2013; Di Paolo & al. 2018; Thibault 2011). "Languaging", on the other hand, suggests that language is a dynamic and metastable (it tends to stability, but it is always subject to change) process of contextual "wording". In this work however, I will only use the neutral "language", except when referring to Thibault's notions of first-order and second-order "languaging". I hypothesize that RP operates a shift from second-order languaging to first order languaging. In this view, a speaker who makes use of RP explicitly highlights a need to synchronize with the listener along a common topic or viewpoint, as a "thematic" or "attitudinal" reorganization. RP links speaker and listener, without exclusively focusing on one of them (it is neither an "expressive" nor a "conative" device).

In the next chapter, I will first introduce RS/CS and quotatives, explain in which regard and conditions they signal that the speech is enacted, and relate a specific branch of RS/CS to RE by subsuming both of them under the principle of "inference constraint".

2. Reported/Constructed speech and Quotatives

RS/CS is a common rhetorical strategy found in various areas ranging from narratives (Moore 1993), to political oratory (Parmentier 1993), religious and esoteric discourse (Janowitz 1993; Wooffitt 2007), testimonies in court (Galatolo 2007), and, of course, in everyday conversations (Tannen 1989; Clift 2007). I will not survey the benefits of RS in all these domains, but will sample through a global survey the characteristics of RS that may lead to a better insight into why and when it is enacted. In 2.2, quotatives will be in turn investigated. They play a crucial role in that they define a RS by isolating it from its cotext and determining its type (constructed/reported). Next, functional equivalences between RS and RE will be brought under scrutiny, in order to bring out converging lines between them and prove that both of them should be addressed together to be better understood. In this chapter, "inference constraint" (that has already been spoken about in the introduction, for exemplary purposes only, cf. *No No* and *I warn you*) will be a guiding principle; also, in this chapter, the notion of "enactment" will remain very close to the one of "animation".

2.1. Reported speech

RS is traditionally divided into indirect reported speech (IRS) and direct reported speech (DRS).

- IRS: Albert said that he was sorry
- DRS: Albert said: "I'm sorry"

Both can combine to form "hybrid" types of RS (Wilson 2000, p.231):

- According to John, I am "neglecting" my job. (mixed direct and indirect quotation).
- John was pretty rude to me. I am neglecting my job!

Banfield (1982) argued that DRS combines a communicative function with an expressive one, whereas IRS is only communicative. In other words, while IRS describes what has been said and how, DRS depicts the RS "iconically". It does not mean, however, that a direct report is more faithful to its source than an indirect one. In fact, a DRS rarely renders accurately a former locution (Holt and Clift 2007). DRS is namely not only used to replay one's own speech or the speech of another, but also to enable the speaker to convey at the same time his attitude towards the reported

utterance (*ibid.*), by e.g. rising the pitch of his voice to show how irritating the original author of the utterance was, what inevitably alters the RS. Also, the quote is textually related to the reporting speech, e.g. Speaker 1: "What did you tell him about the competition you took part in?, Speaker 2: "Well, I said: it turned really bad for me". "It" is a pronoun thematizing "competition" within the progression of the reporting speech. The authentic RS sounded maybe like: "The competition turned bad for me".

Next, Labov (1972) claimed that DRS allows the recipient to "internally evaluate" the point of a story: the direct representation of the reported utterance conveys every piece of information he needs to draw his own conclusions about the recounted event. The speaker does not further need to explicitly evaluate it. Lucy (1993) explains the conditions which orient the selection of either mode:

> Although both modes use verbs of speaking (*verbum dicendi*) to frame and report speech events, the direct form *imitates* or presents the reported speech event from the perspective of the reported speech situation whereas the indirect form analyzes or interprets the event from the perspective of the current reporting event. […] Because of their capacity to instantiate directly the expressive character of language, the direct reports are often seen as more vivid and authoritative. Because of their capacity the describe explicitly the reporter's understanding of the original event (e.g., relevant motives and intentions), indirect reports often signal more clearly the reason for reporting the speech. (p. 18-19)

A further advantage of DRS is explained by the "reduced personal responsibility" claimed by Goffmann (1981). By reporting directly, the speaker changes the footing of the conversation: from the status of "author", he becomes the "animator" of the speech he holds, i.e. he gives voice to words he did not come up with in the first place, and he is consequently not responsible for them. DRS gives therefore more freedom to break good manners or to ruin the seriousness of a conversation without hurting the interlocutor.

The gap between DRS and IRS is not always clear-cut: in spoken French, the indirect structure tends even more to be replaced by the direct one, e.g. "Il m'a demandé est-ce que je viens", "Racontez-nous qu'est-ce qui s'est passé" (pronounced with the same intonation contour as in a regular IRS). It may be speculated that this is due to the massive use of DRS in familiar register because, as we have just seen, its depictive power increases the vividness and authoritativeness of the quote, what guarantees an increased involvement of the listener. This tendency is besides not so surprising, as some languages already merge grammatically D- and IRS, making only in written speech a formal distinction with the help of quotation marks (e.g. in Finnish, indirect and direct questions are the same: "Do you know if she comes?" translated becomes "Tiedätkö, tuleekö hän?, literally "Do you know, does she come?, my own example).

In reality, a RS may reach out to several turns without being signaled each time by a pronoun-plus-speech verb. In this scenario, people do not only report, but enact collaboratively other people or themselves on other occasions (Clift 2007). In a reinforcement of DRS, both participants stage a

(very often) hypothetical scenario which lasts over several Turn-Construction Units[3] (and not only one as in the case of DRS) and extends most of the time a joke initiation. By affiliating with the joke initiator, other participants comply to the shift of footing (i.e. they become "animators" instead of "authors"). Changes in voice quality or prosody (e.g. a falsetto voice), regional accents, or laugh, function as metalinguistic signals which compensate for the lack of report markers and make clear that the conversation is a play. The example below is drawn from Holt (2007), p.72 (for the transcription conventions, see Appendix.)

(1)

(J has bought extra Christmas presents for the children. She proposes to prevent her husband from finding out by paying the bill when it comes.)
```
 1  J: But I thought well I'll go ahea:d,
 2     and, .hh pay for it when it comes
 3     and °he'll never kno:w°,=
 4  L: =°Yeh,°=
 5  J: =(we, [got anything]heh-heh-huh=
 6  L:        [hheh huhehhuh]
 7  L: =[°uhhhh .uhhhhhhh]hh[hhh°
 8  J: =[huh e-huh huh huh]    [.hhehh
 9  J: Ex[cept when Christmas co[:mes a-a-]and=
10  L:   [°°Oh°°                [Y e a h h]
11  J: =.hhhh he says where'd you get all
12     thahheh heh [hn huh] huh=
13  L:             [mehheh]
14  J: =hu [h huh °huh°]°hn°
15  L:     [ .h h h h h ]Santa Claus.=
16     =hhheh-h[eh
17  J           [.hh ↑Santa Claus brought it
18     (in his sle::d).=
19  J: =hn[hih [hn-hn-[hen huh=
20  L:    [ Ye:[ah     [.hh
21  L: = Uh: :[:m
22  J:        [.hhhhhehhhh  °(  [  ).°
23  L:                          [I found a
24        recipe: that I'm gonna try:,
```

The line numbers in bold highlight the enacted scene: it begins with the pronoun-plus-speech-verb "he says", followed by the constructed (= hypothetical) reported utterance "where'd you get all thahheh" and the possible reply imagined by L, pronounced with a rise in intonation: "Santa Claus brought it". This conversation is enacted in several ways: first, and above all, because it is contextualized. Saying that all DRS are enacted because they are analogic is misleading: one must also check if the DRS has achieved its perlocutory effect, and found a resonance in the listener (a joke is a joke only if the partner realizes that it is a joke, and responds adequately to it, by e.g.

---

3   A TCU is the minimal segment of speech in conversation analysis. To be a TCU, a stretch of words must be complete intonationally (i.e. follow a homogeneous intonational bow), syntactically (i.e. unfold cohesively with a coherent grammatical structure) and pragmatically (i.e. achieve a discursive effect).

laughing). Without a context, the status of the RS (enacted or not) is undecidable. In the discussion above, both participants are clearly engaged in a shared scenario that they collaboratively develop. The imagined situation is situated in the there and then, while remaining connected to the deictical space (they make *as if,* but the setting is the same as the one of the reporting situation. It is not a theater set with the adequate scenery, e.g. a Christmas tree and a heap of presents, but the real communication situation). The collaborative, enacted event is also constantly emerging as J and L successively build up on each other's suggestions through active interpretations: "where'd you get all that?" - "Santa Claus" (cooperative answer following up the initiation of the enactment)- "Santa Claus bought it in in sled" (cooperative completion of the answer). Each idea is "added on" the imaginary scene, that both J and L co-construct in a playful manner. A cue to know if interpretation is emerging is to wonder whether utterances could be extended by the filler *you know [what I mean]?*, e.g. *he'd say, you know, like: where'd you get all that…"*. By answering, the interlocutor asserts that he "knows" pretty well what is meant, and the enactment can goes on like this before one of them decides to stop it (the "Yeah" of L. on line 20). Lastly, expressive meaning is thematized: the enactment is introduced by the pronoun-plus-speech verb "he says", before J playfully animates the "he". When playing the role of her husband, she always makes a clear reference to who is speaking through her, what is clearly demonstrated by the laughter. Indeed, what is funny in this situation is the constant endorsing of an other's attitude.

A representation of a scene can also be undertaken by one single person, e.g. "And I answered, Who art thou, Lord?" And he said unto me, "I am Jesus of Nazareth, whom thou persecutest." (Acts 22:8). This time, the nature of the first quote (a question) logically demands to be completed by an answer. In this case, quote markers separate the subsequent turns of "I" and "he" uttered by the same person. On the contrary to (1), where successive TCUs in the speech event correspond analogically to successive TCUs in the enacted event, here a single TCU in the speech event embeds several TCUs in the enacted event. Would it be possible too that one single person enacts a scene over several TCUs in the speech event? Here is an extract from Clift (2007), p.132 for the dialogue and p. 122 for the introduction (again, see Appendix):

(2)

(Marsha, who lives in southern California, is separated from Tony, who lives in northern California. They are discussing the travel arrangements of their son Joey, who is traveling from mother to father. Joey has had the top of his car stolen, and so, instead of driving from southern to northern California, is flying. Tony broaches the issue of how Joey will retrieve his car.)

```
35 Tony:     W't's 'e g'nna do go down en pick it up later? Er
36           somethin like (    ) [well that's aw]:ful
```

13

```
37 Marsha:                              [H  i s  friend]
38 Marsha:   Yeh h[is   friend Stee-          ]
39 Tony:           [That really makes] me ma:d,
40             (0.2)
41 Marsha:   .hhh Oh it's disgusti[ng ez a matter a'f]a:ct
42 Tony:                          [P o o r  J o e y,]
43 Marsha: I- I, I told my ki:ds. who do this: down  et  the  Drug
44           Coalition ah want th'to:p back.h {.hhhhhhhhh/(1.0)}
45           SEND OUT the WO:RD .hhh hnh
46           (0.2)
47 Tony:     Yeah.
48 Marsha: .hhh Bu:t u-hu:ghh friend Steve en Brian er driving
49 up…
```

Usually, a quotation uttered by the same person does not extend beyond one TCU (Couper-Kuhlen 2007)[4]. However, the RS spread here over two TCUs, "I want the top back" and "send out the word". The short inspiration after the first one is actually designed to let Tony take the floor, what he decides not to do. In compensation, the subsequent TCU is intended, on the one hand, to upgrade the first one in an attempt to prompt an answer from Tony and, on the other hand, to secure an uptake that is not coming (Clift 2007).

To summarize: while indirect quotation tends to be predicational due to its referential and characterizing orientation, direct quotation, by contrast, is (re)presentational (Lucy 1993), i.e. it displays all aspects of the original quote (including prosody). The polarity between DRS and IRS is also observed in written speech: neuroimagery has proven that silent reading of direct versus indirect speech activates differently the voice-selective areas of the auditory cortex,

> This suggests that readers are indeed more likely to engage in perceptual simulations (or spontaneous imagery) of the reported speaker's voice when reading direct speech as opposed to meaning-equivalent indirect speech statements as part of a more vivid representation of the former. (Yao & al. 2011, p.3146)

A full-fledged enactment radicalizes a DRS in that it engages all participants in the staging of a hypothetical scenario. This representation actually overrides the original footing of the interaction by transforming the authors of the speech in animators. A single speaker can also represent, in a same TCU, a situation by incarnating two or more persons over several turns in the reported event. But, if a quotation extends beyond one TCU in the reporting speech, then the additional one is an unplanned addition intended to make up for the absence of response from the interlocutor, and to incite him to reply. Normally, a speaker never extends the enacted speech to more than one TCU. This applies also to RE (that I address here in anticipation of further discussion): if I see someone losing his wallet in the street, I may pick it up and shout: "Mister! Mister!". I utter in this case two

---

4   "Not only do quotative marker and quotation together typically form a single TCU; fragments of reported speech
     and thought are overwhelmingly *no longer than* one TCU in non-narrative contexts". (p.94)

different TCUs, because the man, who did not hear my call, consequently did not answer. "Mister! Mister!" is therefore not a true RE (at least not in the sense taken in this work), but an insistence. A second example: if I have a guest who does not dare to eat pre-dinner nibbles, I will insist by saying *"Eat! Eat!"* while reaching him out the dish. In German, we could translate it in *Iss doch!*, in French *Mange donc!*. *Doch* and *donc* are modal particles signaling a reaction to someone else's behavior or words. It presupposes that between the two *Eat,* the partner did not take his turn and that I react to it. Insisting aims thus at triggering a reaction that is not coming, and extends, by definition, over several TCUs. Insisting RE cannot be related whatsoever to speech enactments. Hence, I will not deal with them in the course of this study.

Now that the notion of enacted RS has been more accurately defined, it might be worth expanding on the notion of *perspectivization*. An enacted RS merges the *proximity zone* (the real-time conversation, in the here and now) with the *distant* one (the animated situation, in the there and then). This is made by dovetailing present and past (or non-present, if one wants to include future and conditional) talk. The reported event is always superimposed (spatially, temporally, thematically) on the reporting one. Research on the subject has been conducted by Couper-Kuhlen (2007):

> At the same time reporting past or habitual speech and thought in a non-narrative context can be a way of proposing an assessment or account in the here and now. This merging of the two worlds is facilitated by the fact that non-narrative quoted material is often prosodically and paralinguistically expressive and fades off syntactically into following *hic-et-nunc* talk. Recipients can be observed to orient to the quoted material in next turn *as if* it were an assessment or account in the present. This leads to a prolongation of the sequence, which the quoter – by quoting – initiates. Reenacted assessments and accounts are thus partly there and partly here. In a curious mixture of past and present, they present assessments and accounts as motivated by the past while at the same time opening up this past for renegotiation in the present. (p.119)

Couper-Kuhlen exemplifies with stretches of co-constructed dialogue, where reported utterances are completed by the hearer as if they were situated simultaneously in the here-and-now and the there-and-then. For instance, in conversation (1) drawn from Holt, J constructs the words her husband would have said in the hypothetical reported situation, while L reacts to them as if the reported situation unfolded *during* the reporting situation. The "mixture of present and past" means that the whole reporting situation toggles "somewhere else", thereby affecting the relation between speakers. Indeed, an enacted RS is never an informative report of what happened in the past, but a methodical constraint on the interpretative effort. How can this claim be buttressed? In fact, any RS could be considered as a case of "intended iconicity" (a construction looking like a faithful report)

rather than iconicity alone: what is conveyed is *the assumption* that the speaker imitates a hypothetical or a passed situation. I borrow the term "intended iconicity" from Alvarez-Caccamo (1994) who studied the constructions of sociolinguistic reality through the use of RS in Galician institutions in the 1980s, where the use of standard Spanish was not well regarded and socially connoted. According to her, "intended iconicity" is "a property of speech by which the narration of an event *can count* as isomorphic with the model speech-event" (p.52) In her corpus, apparent isomorphic (= iconic) RS was the object of parodic stylization, while non-isomorphic RS was manipulated by i.a code-switching, e.g. citing in standard Spanish a Galician opponent to discredit him. Shared linguistic ideologies enable an intended verisimilar RS to be taken and interpreted just as it is represented.

> "Code displacement [= "*non-congruent, non-isomorphic* attribution of code choices to a given character under the guise of a faithful reenactment" p.42] represents the reporter's attempt to *construct a(n ideologized) possible world where* the *propositions and its implicated meanings may be true"* […] The relative impunity of reflexive code displacement rests, precisely, on an appeal to shared linguistic ideologies and sociolinguistic knowledge that sustain interpretation." (p.55)

On her account, listeners internally evaluate RS on the basis of implicatures (assignments of discursive relevance) relying on shared knowledge about sociolinguistic facts.

I would like to extent this assessment to enacted RS with the help of Sperber and Wilson's *relevance theory*. In their seminal book published in 1986, Sperber and Wilson claimed that human communication heavily relies on inferences. Textual and contextual cues prompt an audience towards the communicated meaning. Because human cognition is geared towards the maximization of relevance, the message is interpreted against the contextual backdrop that makes it the most consistent. The more a message has positive contextual effects (i.e. the useful results of the message) and the processing effort is low, the more it is relevant. They also argued in their book that the cognitive mutual environment (the ideas shared with somebody else in the interaction) is not based on *mutual knowledge*, as assumed by Grice, but *manifestness*, defined in the following way:

> An assumption is manifest to an individual at a given time if he is capable at that time of mentally representing it and accepting its representation as true or probably true (Wilson 2000, p.421)

An assumption (that is "thoughts treated by the individual as representations of the actual world" (Sperber & Wilson 1995, p. 2) "can be manifest to an individual if it is merely capable of being non-demonstratively inferred" (Wilson 2000, p.422). Thus, a speaker only needs to ostensibly makes manifest to an audience that he intends to make manifest a set of assumptions. A RS ensures that both the speaker and the audience share a same set of assumptions by making an utterance mutually manifest. The faithfulness of the RS does not matter; what matters is that it enables people to *share similar dispositions* towards what is said. The form of the reported utterance (i.e the language used – Spanish or Galician – or the prosody – falsetto voice, depth or intensity) will then

suggest how the RS is to be interpreted. The hearer, through a string of inferences, triggers one or many possible interpretations of the quoter's meaning, by order of accessibility.

I will now try to put it more simply: "manifesting explicitly an intention" in the message leads to what I qualified in the introduction of "inference constraint". In some cases, speech enactments maximally constrain inferences about what should be understood (i.e. a clarification of intention). This is most obvious when the RS is framed by a performative that presents (and not represents) the speech act, in other words: when the utterance is not only produced, bu also put into perspective by a meta-verb that specifies which speech act is performed. Let's take typical examples: X and Y are brother and sister, X becomes one day angry and says to his sister: *I order you: "Don't do this ever again!"*. The explicature (i.e. what is explicitly pointed out, in contrast to implicature) *I order you* is supposed to steer Y to the right inference: "You may believe that, as your brother, I can't order you around and that I must be joking, but don't be mistaken: I'm really giving you an order". It works too for RE (that I here too address in anticipation): X receives a cheese grater as a gift for his birthday, and says, a bit confused, *Oh, thank you, thank you very much*. *Thank you* is repeated by X to let believe that he is happy, while actually he is not, but does not want the others to infer that he is disappointed, because it is impolite. A second example: X's mother says to her son: "Will you please tidy your room?", and X answers, listless: "Yes, yes". This time, X has clearly not the intention to tidy his room, but he wants all the same his mother to believe that he will do it (or maybe well, and he just wants to defeat the assumption presupposed by the question that he may not do it). In any of these cases, a speech enactment has to do with the handling of other's inferences. As a result, a RS is never a mere informative report of what happened in the past, but a strategy deployed to have concrete effects in the reporting situation.

Sometimes, this strategy is unconscious and only aims at creating social ties: our daily speech is deeply intertextual, grounded on "hidden quotations" (Fonagy 1986), where the *apparent* split attitude between animator and author is minimized. Code-switching is a typical case of such reflexive use: it implies namely that the speech is not only produced, but animated. Switching from a standard language to a marked one (e.g. a dialect, a slang) creates a feeling of intimacy, humor, and sense of belonging proper to the use of enacted RS (see Mertz 1993, where the use of Gaelic is characterized in such words by interviewed Irishmen). Reported utterances take sometimes the shape of usual utterances. Irony, for instance, is a superimposed voice that conveys an attitude at odds with the expected one:

(1) Looking at her son's messy room, Mom says, "Wow, you could win an award for cleanliness!"

17

(2) On the way to school, the school bus gets a flat tire and the bus driver says: "Excellent! This day couldn't start off any better![5]

Even in the absence of RS/CS markers, there is evidence that the bus driver and the mother report what they could have said in *a possible world* where their utterances would have been true, with the intention of making a contrast between this possible world and the actual facts. Setting the present situation in an animated one brings a bit of humor to a potential tinderbox (if the mother had said: "your room is a true mess!", it could have led to an argument between her and her son), and creates a moment of complicity and involvement, instead of bringing about a conflict. Being ironic is involving because saying the contrary of what one has in mind or what is implicated by the context encourages emerging interpretation: the sentence "Wow, you could win an award for cleanliness", could be extended with the little footnote *you know?* Or preferably just a wink indicating that active interpretation is needed, while "your room is awfully dirty" not. But once again, in the case the child does not grasp the foregrounded expressive meaning of his mother's utterance, no perlocutory effect, and subsequently no speech enactment, is made possible. Finally, let us note that irony is also reflexively used (cf. the bus driver) to cool oneself down in a stressful moment, such as it happens when a tire goes flat. Irony, as a speech enactment, affects thus the relationship between conversations partners, even when they are embodied by the same person.

What should be remembered from this section is that a direct report, under given conditions, modulates the way speakers interpret and relate to each other. It thematizes an expressive meaning, and, by doing so, prevents, fosters, guarantees or subdues possible inferences. Due to the fact that enacted RS, or animated speech in a very broad sense (which includes "hidden quotations": code-switching, irony) manifests a specific expressive meaning, a few leeway is given to other possible interpretations. Moreover, a high degree of affiliation is requested for an enacted/animated speech to be successful, i.e participants must "play the game" and accept to enter in the virtual world that is proposed.

In the meantime, I've also drawn some parallels with echoes (*yes yes, thank you thank you, no no*). A first distinction has already been made between insisting RE which extends over two TCUs or more (*Mister! Mister!*) and RE which unfolds in a single TCU. In point 2.3, equivalences between RS and RE will be systematically made to legitimate their joint study.

Until now, only RS in itself has been discussed. There is still one important aspect of RS that I so far deliberately left out: quotatives, the "hinges" or "signals" that often frame a speech enactment

---

5 https://www.softschools.com/examples/grammar/verbal_irony_examples/122/ (accessed on February 12 2020).

18

and single it out from its propositional environment. Their significance is underscored by the fact that they are the necessary *and* sufficient condition to enact a speech: namely, once used as fillers, they do not require a following RS anymore, yet do not lose their role as pivot between digital and analogic speech. After having discussed enacted RS in the core, it is now time to devote a new section to the switches triggering it.

2.2. Quotatives in RS

Quotatives are selected according to the type of RS they frame. *Say that X* and other speech verbs-plus-conjunction are referential and focus on the propositional content only (Güldemann 2001), while quotatives without subordinate clauses (*say: "..."*) are associated with analogic RS (they are supposed to report exactly what has been said). They also specify the *mood* of the RS, divided into *realis*, *hypothetical,* and *situational*. The first section will focus on the new quotative "like", which is of particular interest. In the next one, I will try to find out whether modality, by steering the interpretation, is responsible for the enactment of the speech. Lastly, quotatives and fillers will be paralleled. Fillers, namely, have the ability to put a whole portion of the speech in quotation marks, not only a determined section.

2.2.1. The new quotative "like", between reported speech and reported thought.

The selection of a quotative results from a convergence of several axes: type, person, and mood. For example, *He said: "…"* is direct (type), 3P.S (person), and *realis* (mood), because it reports an authentic utterance made by a *he* in the past. But in real life, these categories are often fuzzy: first of all, quotatives do not necessarily predict the type of the ensuing RS. In Finnish, *sanoa että X,* "say that X", initiates a IRS as well as a DRS ("although *et(tä)* has been considered often as a marker of indirect reporting, in spoken language it is quite commonly used with direct reported speech as well", Haakana 2007, p.157). The Dutch equivalent of "like", *van,* introduces sometimes indirect speech (*van* not only combines with direct speech, but also occurs regularly in combination with indirect speech", Coppen and Foolen 2012, p.259). In spoken french, *dire que* sometimes introduces direct speech, the most common example being *Il m'a dit que oui/non* (instead of the direct equivalent: *Il m'a dit: "oui/non"*):

19

(3)

Bielefeld _ situations de contact – cadre universitaire ~ Cours de figura

I   il a pas il a exu- existé après oui (.) j` `ui ai j::- j'ai
téléphoné à doris après pour demander [si] ça s'était arrangé
(.)
S   [oui]
I   elle m'a dit que non il fallait que j'envoie un papier j'ai envoyé un
papier voilà ça marchait
(0.2)
S   hm hm

(.)   = micro-pause (less than 0.2ms)
(0.X)= length of pause
[ ]   = overlapping turns*
:::   = sound stretch
-    = word initiation
`   = nonstandard elision
Red = significant part
*Due to technical issues on CLAP, overlaps are not aligned

"I" uses a verb-plus-conjunction because she first intended to say "elle m'a dit qu'il fallait", but inserts ultimately a "non" in DRS to express the denial. Doris' voice is thus momentarily represented by the *non*, before the IRS takes finally over.

Below, another example of co-occurring DRS and IRS:

(4)

Histoires racontées par des enfants - (gre)noble

+ = phrase boundary

E   donc tu crois que dans la vie il faut être un peu méchant
(question)
NIC non + j'ai pas dit ça
E   et qu'est-ce que tu as dit alors
NIC j'ai dit que ben ben si si on est gentil d'accord mais si on est trop
    gentil aïe aïe aïe
E   aïe aïe aïe pour nous ou pour les autres
(question)

NIC says *dit que* under the influence of the *qu'est-ce que* in the previous turn. Cotextual factors have thus also an impact in the use of a quotative-plus-conjunction combined with a DRS. On closer inspection, it can be noticed that NIC does not actually report what he *said*, but what he *meant*, without interrupting the flow of the conversation (*si si […] d'accord* is a direct reaction to E's reformulation of NIC's words). It is an enacted RS, because the expressive meaning is foregrounded and the inference is constrained ("I did not exactly say this, but that"). A RS which constrains an inference is usually a performative in the present tense, but in (4), the quotative is in the past tense; paradoxically, it also presents the speech: *I have said: "…"* should automatically be followed by a representation of what has been said before, but here the quotative is only a means for NIC to present his reply to the clarification request of E. In fact, it would be grammatically correct to substitute *I say* for *I have said,* but the previous utterances, which include each a quotative in the past ("J'<u>ai</u> pas <u>dit</u> ça – et qu'est-ce qu tu <u>as dit</u> alors?) prevent any abrupt shift to the present (and consequently to a performative). The issue confronting NIC is that he wants to present his speech

by framing it with a performative in order to constrain the inference, but he cannot due to the conditioning of the previous turns (i.e. the *I* and *you have said*). This disruption manifests in NIC's hesitation (*ben ben*). By the way, he also utters a following, concessive *si si*, which is a second enactment (performed this time by way of a RE) pointing out that NIC does not want let E infer, on basis of his denial ("J'ai pas dit ça") that E's interpretation ("tu crois que dans la vie il faut être un peu méchant") is completely wrong.

In the previous paragraph, I gave a concrete example of a "true" quotative (not only a performative, such as *I warn you* or *I order you*) that, in specific conditions, does not so much report a speech as clarifies its meaning. Similarly, the "new" quotative *like* does not report a speech, but an expressive attitude clarifying how something must be understood. A typical example would be somebody who narrates an event that happened to him in the past, and, on a given point, would say "And then I was like: oh my god!" to describe a feeling of surprise he experienced. In certain conditions, *like* can be replaced by a reported thought (a reported thought is a kind of reported speech, where the *I said: "..."* becomes *I thought: "...",* see fragment 5 below for an example*)*. A reported thought has the ability to intensify the internal evaluation when a mere reporting of a situation does not help the hearer to assess it. When retelling an event, a speaker may want to intersperse his story with the depiction of how he reacted to it, so that the listener can align to the displayed attitude. It guarantees the listener's involvement in the story:

> With reported thought the narrator can more clearly show how he or she evaluated the (verbal) actions in the situation depicted. In effect, reported thought offers more explicit evaluation than reported speech but nevertheless situates the evaluation in the story-world, not in the present interaction. (Haakana 2007, p.160)

Reported thoughts, often found in complaint stories, are, among others, a way for the speaker to make "a picture of the situation in which the speaker possibly resisted the antagonist by criticizing him or her" (Haakana 2007, p.167), i.e. to show oneself in one's advantage during an actually unfavorable situation that one retells afterwards to others (the fragment below is originally in Finnish, but I reproduce hereunder only the translation in English; see Appendix for conventions):

(5)

```
18 L:  .mhh >then he said to me @Well, (0.6) such
19     questions are so difficult to make, h@
20 J:   M(h)m[: m (h)h        ]
21 L:         [.mt I thought that @<fi:ne fi:ne>.@
22 J: =mh [h
23 L:      [he could prac£tise making them a bit.£
(Haakana 2007, p. 162)
```

In this scene, a student (Leena) reports to her friend Jari the answer her professor gave to a complaint she voiced to him. His reply apparently did not not satisfy her, and, when Leena describes later on her reaction to it, she uses a reported thought (L.21: "fine fine") that represents her inner attitude at this moment, and suggests that she passively resisted to the unconvincing answer of her professor. When producing *fine fine* (Finnish "kiva kiva"), her voice is clearly animated: this is showed by the irony (actually, the situation is anything but fine). The direct reported thought enables Leena to enact her speech in order to seek the affiliation of Jari. In my view, the verb of thought is interchangeable with the quotative *like* ("niinku" in Finnish): "I was like: fine fine".

Does it mean that the new quotative and the quotative of thought are similar? An utterance such as "I was like: what?" evokes the surprise of the speaker in a past situation. It is a private feeling which can possibly be reported thanks to a quotative of thought (because thought and feeling are both inner states). At the same time, a feeling can be expressed overtly: indeed, what follows *like* can be showed non-verbally: *and I was like: [frowning/grinning/sighing/…]*. Thus, two apparently opposed categories are conflated in it: on the one hand, it is a private feeling or attitude (viz. rather a thought or emotion than real words), but on the other hand something that can be deciphered through verbal or non-verbal cues in the reporting event. This paradox could be solved if we think that the feeling conveyed by *like* does not preexist to its non-verbal or verbal representation; the feeling only begins to exist through the enactment that shares this very feeling: the motivation of a *"like + RS"* is not to *share* an attitude previously experienced, but to *generate* an attitude through a strategy of perspective alignment (= enactment) designed for argumentative purposes. This can be illustrated by the following fragment (Buchstaller 2002, p.7):

(6) *Cooking*

B: so I enjoy you know cooking thinks to take over to her hou[se or-
A: [oh that is nice,
B: yeah and it is fun for me to do that,
It is something I enjoy doing,
It is funny though **it's like** 'I don't really want to cook for us' @ [@
A: [jeh @@

It is impossible to report words that have not been said, what is more that pertain to emotions rather than verbalized thought (*what!*: surprise, *it's terrible*: dismay, *I don't really want to cook for us*: reluctance,…). These feelings have nonetheless to be directly reported so as to be co-experienced and sanctioned by the listener. As in all enactments, if the foregrounded expressive meaning is not grasped by the listener, it is invalidated (again: a joke, if not recognized as a joke, is not a joke). In the same way, if the listener does not understand that "its like I don't really to cook for us" conveys a particular attitude (its expressive meaning), then the feeling is not shared, and

thus invalidated. In this conversation, A validates with a "jeh". The enactment is successful. *Like* is different from *I think* because it introduces a constructed speech that actively involves the listener (i.e. the listener must play the game, he guarantees through his positive reaction that the speech is enacted), while a reported thought does not necessarily expects the listener to enter in the displaced context that is proposed.

Actually, the quotative *like* thwarts any failure of the enactment, because it denotes approximation:

> The content of the quotation can only be an approximative rendering of the whole emotional and contextual situation. Using like with its approximative-comparative semantics signals the possible non-equivalence of what is reported and the actual utterance. The speaker retains a reduced responsibility with respect to what was said and how, as a quote introductory like does not commit her to the form and the content of the quote. Like then functions as a hedge, both on the referential-epistemic, as well as on the interpersonal-pragmatic level. (Buchstaller 2002, p.4)

By denoting approximation, *like* signals that the speech is constructed: in general, many cases of RS have never been uttered nor thought before, but are made up in a current conversation. The speaker using *like* communicates explicitly that he wants to share an experience or a point of view in the actual interaction. When introduced by *like*, a RS is always constructed on the very moment of the speech event, even if they are supposed to represent something. Here is an example from Buchstaller 2012, p.5:

(7) *using plastic grocery bags as lunch-bags*

B: Yeah in fact I have one today
A: rig[ht.
B: [the only problem with those is sometimes they got holes in the bottom
A: yeah   [they
B: [and @@ **it's like** 'whoops there goes my chips
A: [yeah
B: okay fine'.
A: uh uh

@ = laughter
_ = primary or main accent
[…] = overlap
'…' = signals for start and end of quote
Bold = significant part

'Whoops there goes my chips' is a verbalization of an *act;* it represents verbally the gesture of somebody who tries to catch his falling chips. This type of RS goes most of the time with gestures or facial expressions of the speaker who physically enacts, through a gesture which situates something in a displaced deictical space, the scene described. In a typical enacted CS as: "I was like: what?", surprise is expressed through a recognizable behavior which is not *felt*, but *performed*: the speaker is not "surprised", but, as it were, is "doing surprise". Imagination and play are thus necessary to put forth such an enacted performance, which ultimately aims at making the listener co-experience the emotion, or perception, as in the following extract:

23

(8) Corpus : **Apéritif entre ami(e)s - chat** (**15c/4a6**)

JEA (../..) limite jura saone et loire ain tu vois c'est
JUL hm hm\ bah saint amour en fait ça ça m` dit rien/ mais moi j'étais
      dans l` bas jur""a: pendant les vacances
JEA pis dans quel euh: `fin t'as p`t'être [un nom] d` village tout ça euh
JUL   [pf .h]
JUL ouais c'était un p`tit village [en fait c'était genre] montagneux `fin
      c'était [pas   ] grand chose quoi l` jura euh:
JEA   [c': c'est quoi/   ]

CLA
      [< ((rit)) uhm>]
((rires))
JEA ouais dans l` doubs ouais (.) nan mais ouais le nom du village tu:
JEA [et euh ouais non
JUL [nan j` m'en rappelle plus le nom hm hm
JUL c'était vraiment petit et puis euh: .h mais euh j` me rappelle encore
      que c' était joli (../..)

JUL has forgotten the name of the village in which she spent her holiday. What she therefore (desperately) tries to do is making her friends visualize the scene. When she qualifies the region in which the village is nested as "mountainous", she shares an *impression* (it looks like an alpine village). The adjective "mountainous" is not descriptive, but depictive, as it is supposed to reflect a perception, which, in case of success, the others are able to co-experience visually. In the example above, JUL fails, and, upon request of telling the name of the village, keeps saying that it was "little" and "nice". In the video of the aperitif, JUL accompanies *genre* with a concomitant gesture, while she sat with the arms crossed the minute before. The gesture in question is an open palm inciting the listener to watch the scene depicted by imagining it, and not only to decode her patchy explanation:

I interpret this gesture as a prompt to arouse imagination. In French, the equivalent of "you know" is the filler "tu vois?". "Tu vois?" indicates that what precedes must be visualized (as does the gesture above). For example, I had recently a conversation with somebody who said: "Je suis arrivé vingt minutes plus tard – vingt minutes, tu vois – et il m'a dit qu'à cause de moi, il n'a pas pu

allé faire les courses". Not only the first part of the utterance is repeated, but the filler is added on the RE. If I had taken the sentence literally, I may not have guessed what was the communicative intention of the speaker (i.e. that twenty minutes is a ridiculously amount of time and that the speaker could not help for the delay.). Much of what is to be understood was not included in the propositional content, therefore I have been incited to "see" the situation analogically, to re-live it by myself in order to empathize with the interlocutor. In the same way, everything said by JUL that lies before the pivotal word *genre* is only a symbolic exchange of information, everything after is a multimodal (i.e. a combination of words and bodily movements) co-experiencing of the information.

In a distributed-enacted view, *presentational* corresponds to *first order languaging*, and *propositional* to *second-order languaging* (cf. introduction), as Thibault (2011) puts it:

> [First-order language] is grounded in the intrinsic expressivity and interactivity of human bodies-in-interaction. […] The realtime bodily dynamics of interacting agents are, in turn, constrained by second-order patterns emanating from the cultural dynamics of an entire population of interacting agents on longer, slower cultural-historical timescales. Second-order patterns are intrinsically normative. Lexicogrammar is one manifestation of such second-order constraints on first-order languaging dynamics. (p.1-2)

I would like to argue that the quotative *genre* in the extract above initiates a shift between second-order languaging to first-order languaging. *Genre* announces that all participants have to synchronize around co-experienced imaginary scene or performance: somebody saying recurrently *I mean* or *like* tries to project himself in an imaginary performance. In the same manner, the *genre montagneux* of JUL is supposed to turn the conversation into a visual, imaginary scene. In the latter case, the speech is (mainly) expanded by gestures whose goal is to generate in the spirit of the interlocutor a specific picture. New quotatives indicate thus that what follows needs an activation of imagination.

Equivalents of *like* and *genre* in other languages can also help to understand its role as hinge between propositional and representational content: Lucy (1993) studied the particle *ki-* in Yucatec Maya, which corresponds pretty much to English *be like* and *go*. Like them, it can represent a non-referential sound (e.g. The door went like: "iiiiiiii"). The particle *ki-* may replace any other speech verb, and does not specify if the quoted utterance is a speech or a sound. Therefore, *ki-* is a particular word that does not belong to any paradigm of the language, and, what is more, has a narrow morphosyntactic flexibility. It merely signals a following RS or sound. Lucy calls it a "metapragmatic presentational": it reduces the reporting speech to a pure metalinguistic form which does not convey any information, only foregrounding "the specific of the utterance and predicates

as little as possible about its form, content, or function." (p.99), "In this sense, it is the most vivid reporting form - it directly presents the reported event as an "event"(p.118).

Does the form *be like* obeys to the same dynamic? *Be like* is definitively less predicational than *say*, because, as previously said, it does not specify how and whether the RS has been actually said. It only constructs a speech reflecting an attitude. The signification of the quotative *go* is even more unspecified, and, like *ki-*, proportionally more frequent than *be like* to reproduce sounds unrelated to language. What they all have in common is that they isolate depictive parts of the speech from a descriptive environment, thereby prompting us "to imagine what it is like to experience the thing depicted" (Dingemanse 2015, p.954). Simultaneously, speakers tend to highlight what follows this type of quotative (most of the time by means of prosody, e.g. modulating the pitch of the voice) to put them on as a performance. *Ki-* indicates too that what follows is not to be understood, but interpreted ("they invite us to take a particular stance", *ibid.*). Dingemanse's view on quotations is in line with what I call a speech enactment: "Consider quotations: they are often embedded in our utterances, and yet they are at the same time images – depictive reproductions – of other utterances, produced in such a way as to enable the listener to imagine what it is like to experience the thing depicted" (*ibid.*, p.951). Quotatives thus frame and put in the foreground a depictive part of the speech, be it an ideophone (or impersonal echo, as I call it in chapter 3) or a RS/CS. Their role as switchers is thus obvious.

2.2.2. Epistemic modality

Quotatives can also be distributed according to their *mood*, that is "how, in quoting, speakers index their relationship and attitude towards the quote and express the general probability of the occurrence of the quote" (Buchstaller 2012, p.5). Discourse modality is by nature expressive, because it "foregrounds certain ways of interpreting the propositional content in discourse; it directly expresses the speaking self's voice on the basis of which the utterance is to be interpreted. (Maynard 1993, p.38-39). Quotatives are an integral part of RS in the sense that they indicate how to interpret what follows. They disclose the speaker's attitude towards what is said. It raises the question whether quotatives enact the speech precisely because they are modal. In this case, what I called "inference constraint" would only be a sub-type of modality. This question deserves fuller exploration.

Comrie (1996) and Akatsuka (1986) have proposed to classify quotes on an hypotheticality continuum ranging from *factualis* to *counterfactual*. Quotatives are selected according to the mood of the RS they introduce. *Realis* corresponds to faithful reproductions, e.g. He said [exactly] "I

don't want to go"; *hypothetical* to utterances made to express the mental state of a speaker (see the example over plastic bags above). The status as thought or verbal expression is in this case left open (this indeterminacy is called "verbal uncommited thought" in Chafe 1994). *Situational* means that there is no reproduction (it is constructed). The "cooking" fragment illustrates a case of *situational* RS:

Cooking

B: So I enjoy you know cooking things to take over the hou[se or
A: [oh that is nice,
B: yeah and this is fun for me to do that,
    It is something that I enjoy doing,
    It is funny though it's like 'I don't really want to cook for us'@[@
A: [jeh@@

Here, the attitude of the speaker is clad "in the format of a quote" (Buchstaller 2012, p.7). Contrary to the realis and hypothetical mood, there is no representation. B just decides to convey her attitude with a quote rather than with a regular proposition.

The quotatives are not used indistinctly: according to the mood of the quote, one is preferred over the others, as shown in this table from Buchstaller (2012, p.8):

Table 2: Distribution of degrees of hypotheticality (in %) per all quotatives

	realis	hypothetical	situational	habitual
say	52	7	10	18
go	45	17	1	28
like	22	28	14	30
think	0	51	38	7

p= 0.001

Say and *go* are preferred when it comes to realis speech, while *think* is assigned to hypotheticality. The quotatives *go* and *like*, being verbally uncommited, are to be found in the two categories, albeit *go* is used for higher epistemic stances. *Go* and *think* are opposed in the situational category: "go does not introduce situational quotes, it does not have an equation function between a quote and a situation. This suggests that go needs a real communicative situation, if only a hypothetical one." (*ibid.*, p.8). Reported thought, on the contrary, ought not necessarily to render a previous utterance: it can merely put into words an inner mental state situated in the actual moment of the conversation. Eventually, *habitual* is somewhat separate from the others, because it is a rather a complement than an apart category on its own, which is not of much use in this study.

An additional particularity is that *think* introduces less representational content than the others quotatives, as shown in the table below (*ibid.*, p.9)

Table 8: Co-occurrence of quotatives with sound effects

mimesis	%	Number
go	76	132
like	69	159
say	42	50
think	20	21

p<0.001 χ^2: 110.634 (df 3)

Indeed, thoughts, as inward oriented speech, do not naturally go with voice effects, because *to think* implies materially unrealized sounds. *Like,* on the contrary, is indeterminate: the quotes it introduces, even if they have not been produced, are packaged in an ambiguous way; they look alike quasi-direct speech: their status as RT or RS is completely left open.

The problem of modality is that it applies to the utterance only, not on the relation between participants. If one has to classify quotatives, one must take into account how they aim at sharing a point of view (and not at coding a mood). How do "go", "like", "say", "think" let know, each in their own way, what is the "point" (in its broadest sense") of an utterance? Sharing a point of view equates not to lead the partner towards a particular interpretation of an utterance, but to give him information about the *stance* (or perspective) that he has to take.

Actually, quotatives occur often on their own. Their inherent role is not to modulate, because frequently, they do not even apply to a particular stretch of speech, as a RS/CS. In isolation, they become fillers, and their only utility is to ensure that a common point of view is sustained between partners of conversation. In that case, this is not only the portion of text between quotation marks that is enacted, but entire stretches of conversation.

2.2.3. Fillers and quotatives

A less studied dimension of quotatives are *self-quotatives*. e.g. "Regarding to this fact, I answer you:", "I'd say…", etc.. Self-quoting may be done in direct-style (e.g. "I tell you: don't yell at me so rudely), or indirect-style ("I tell you that I didn't know"). They play a decisive role in politeness- and face-theory. For instance, Cohen-Achdut (2019) showed, on basis of articles written by women in Hebrew in eastern Europe in the nineteenth century, that diverse types of constructed utterances introduced by self-quotatives may be used to prevent face-threatening acts and act as hedges against the dominant male community. Cohen-Achdut has put light on three of their functions: (i) construction of past discourse events ("I said"), (ii) construction of discourse events that might occur in the future ("If you ask me, I would answer…"), (iii) discourse events about which it is openly declared that they will not take place ("Far be it from me to say that…").

Quotatives, and above all self-quotatives and quotatives *through simile,* are so widespread in discourse (in particular as hedges, but not only), that they often evolve into fillers, e.g. the intervention of a philosophy student during a lecture: "But if you said like the blanket is comfortable is that a fact like because it depends on like if somebody could think…"[6]. In the philosophy student's remark, it could be assumed that "like" might complete "said" to strengthen the potential modality of the following RS, and the next ones are fillers. The constructed speech initiated by "said like" triggers a potential scenario that primes two additional "likes" that keep signaling the purely approximative or speculative nature of the proposition. In my view, the student tries rather and above all to seek the adhesion of the teacher by not looking assertive: fillers often convey uncertainty and doubt, and are important as discourse markers to stay polite or to avoid any face-threatening act, e.g someone who wants to decline an invitation without offense would say: "umh, er, let's say, well, I don't think I'm in the mood tonight, you know?".

The other well-known filler, *I mean,* has a phatic function that helps to maintain contact between interlocutors. It is also present in the elided *you know [what I mean]*. "You know" introduces often a constructed speech, e.g:

a) you don't just need some good advice or a cute little cliche, because we're really good at Christian cliche**s, you know**,' Uh, let go and let God.' But in case you hadn't noticed, sometimes if you let go, and let God, they will cut off your power.
b) So I guess**, you know**, what I'm saying is,' Who gives a bleep?' "
c) It's not super**, you know**, challenging. It was just kind of like,' Cool, I can see how they would have this.'
(from the corpus of Contemporary American English, https://www.english-corpora.org/coca/)

The "you know" normally signals the speaker's assumption that his interlocutor knows what he means. However, in the examples above, each speaker specifies all the same what he intends to say. In a., "you know" merges with a virtually present quotative ("one would say") and is directly followed by the CS. In b., the two appear consecutively, and in c. "you know" precedes "challenging". This last example shows that "you know" is not always interposed between the emphasized part of the sentence (which can always be set in quotation marks: it's not super, you know, **"**challenging**"**) and the RS, but that it is sometimes inserted before them. In my view, *you know* is supposed to prompt a back-channel from the listener (nodding, …), who is reminded of his duty to punctuate and sustain the actual speech, in order to show that he understands, so that the speaker can initiate a new turn and further develop his ideas. As a filler, however, *you know* is also used reflexively, in which case the speaker does not expect his addressee to react; it is rather a self-stimulation that boosts or reorders (typically) a confused or an emotionally charged speech, when the speaker walks on eggshells and has difficulty in expressing his thoughts. It leads to situations

---

6   https://www.youtube.com/watch?v=S1RPRp5bDgg&t=613s, 10:00-10:10 (accessed on 22 April, 2020).

such as saying *you know* (or *I mean*) even before the onset of the utterance. Each verbal production demands a constant interpretive feed-backing; resorting persistently to fillers is a cue that the interactional feed-backing is impaired, so that the speaker has either to prompt a response from the listener, or to interpret himself reflexively. Both strategies are realized through fillers.

In French, the expression "entre guillemets" (literally, "in quotation marks", but similar to *like*) in e.g."Je lui ai demandé de me rendre le dossier X. Il a refusé tout net. J'étais - entre guillemets - choqué par sa réaction"[7] ("I asked him to give me back the file X. He gave a flat refusal. I was like "shocked" by his reaction), demonstrates more clearly the link between quotative and filler: the quotation marks are explicitly formulated, not to quote, but to mitigate the coarse feeling which is expressed ("choqué") and does not fit well in the reporting situation. The functionally hybrid quotative/filler introduces the blunt words that will follow, but is never part of them: it is therefore subjected to the required language register: "genre" is familiar, "entre guillemets" is neutral, "passez-moi l'expression" formal.

As an illustration, the summary of Pierre Merle's book "De nos tics de langage, panorama aussi raisonné que possible" (2008) is a striking example because it coincidentally talks about the minimizing effect of fillers while permanently making use of quotation marks:

> En fait, le tic de langage envahit quelque part l'ensemble de notre discours. Bien évidemment, il sert à asseoir le propos et c'est vrai qu'il amène effectivement de la proximité, une connivence au niveau de la conversation, *si vous voulez*. Mais reste qu'il est, au jour d'aujourd'hui, ce qu'on pourrait appeler *entre guillemets un "grand indispensable inutile"*.
>
> Car le tic de langage ne sert à rien, sinon à se donner le temps de réfléchir à ce qu'on va dire tout en parlant, à quêter l'adhésion de son interlocuteur, mais aussi à *atténuer, arrondir ou rendre "politiquement correctes"* chacune de nos affirmations.
> Il y a aussi le tic *"tribal"*, qui marque l'appartenance à un groupe : les cours de récré plébiscitent c'est clair (ou juste clair) à la place de "oui", trop pas à la place de "non" et en fait en début de toute phrase. [8]

The first alinea shows an expression similar to "passez-moi l'expression": "si vous voulez", which is reinforced by the concession "c'est vrai que". "Entre guillements" redundantly strengthens "ce qu'on pourrait appeler", and is there to make the apparently contradictory "grand indispensable inutile" acceptable. The author constantly attributes to others the words the he writes himself. "Politiquement correctes" is also put between quotation marks, whose mitigating impact ironically reflects the true meaning of "politically correct" (i.e. reducing the directness of some words to make it acceptable, what is obviously the effect of the quotation marks.). "Atténuer" and "arrondir" are reinforced by and pave the way of the following, less neutral expression "politiquement correctes".

7   https://www.lefigaro.fr/langue-francaise/expressions-francaises/2016/08/04/37003-20160804ARTFIG00071-du-coup-aller-sur-ces-erreurs-de-langage-a-eviter.php (accessed on April 22, 2020)
8   https://livre.fnac.com/a2353583/Pierre-Merle-De-nos-tics-de-langage-panorama-aussi-raisonne-que-possible (accessed on December 13, 2020).

"Tribal" is put into quotation marks because of its negative connotation; they signal that the word has to be taken literally, in the sense of "qui porte l'appartenance à un groupe" (the relative clause, again, facilitates the interpretation by paraphrasing the questionable "tribal"). Each time, the assessments are alleviated by repetitions, quotations marks and attribution of words to others pointing out that a word or expression has not to be taken literally (something has not the be understood like *this*, but like *that*). Our inclination to use all types of "quotation marks" in order to spare our audience may explain why they develop into fillers.

Finally, quotatives are quite often completed by discourse markers, most of the time "well", e.g. "I get offers all the time to play villains in certain things and, at that point, when they said to play Freddy's father, I said, well, as long as I don't have to look like Alice Cooper."[9]; "They kept saying both heartbeats are really strong, and I thought well, that's good 'cause I'm having a baby"[10]. "Well" alert that what will follow is a quote (and thus an animated speech):

> "Discourse markers and other 'expressives' are well-known signs of direct reported speech and thought. In other words, when a turn-initial discourse marker such as 'well' appears […], it may indicate a shift in footing to a reported speaker". (Couper-Kuhlen 2007, p.123).

"Well" is a hinge that "link the upcoming turn to the prior in the face of [an] apparent disjunction" (Clift 2007, p.123). Indeed, the discourse marker links the quotative, which situates the quote in the past, with the quote itself, constructed in the reporting situation. It could be assumed that "well", when used as a filler, also enacts what follows, without pointing to a RS in particular.

In conclusion, fillers constrain inferences as much as any enacted RS does: they indicate that the speaker does not authorize, but animate his words, and consequently does not take responsibility for what is said. Therefore, the listener is asked to not understand literally, but to infer what is meant in a controlled way. What is important is the communicative intention behind the words, or the point of view that is conveyed. The expressive meaning is put in the foreground, and, in case of inference constraint, almost imposed. The listener has no other option than to adhere to the view of his partner, because the latter, by making use of fillers, presupposes that he will do so.

We have seen so far that small words can put the continuous flow of the speech "into brackets". In the next chapter, I will show that doubling a word has the same effect (even if the role of RE exceeds a constraint on inferences, see next chapters).

---

9   https://context.reverso.net/traduction/anglais-francais/Alice+Cooper+play (accessed on December 13, 2020)
10   https://context.reverso.net/traduction/anglais-francais/saying+both+of+them (accessed on April 22, 2020)

2.3 Comparison of RS and RE

In this section, some of the pragmatic roles of RE will be equated with the ones found in RS, and the extent to which the inference constraint principle can be applied to RE will be examined. This cursory comparison will demonstrate that both of them are functionally similar and ought not to be dealt separately. The end objective is, at a later stage, to understand why RS and RE are similar from an enacted point of view.

First, I would like to specify what counts, in my view, as RE and what counts for RD (that I will address later): RD covers conventionalized lexemes (e.g. ding-dong, razzle-dazzle, zig-zag). RE, on the other hand, have a simplex that is always independently meaningful (bye-bye, yes yes, no no, fr. Alors alors?; Ça va ça va; bon bon, très très…) and a reduplicant that never changes the informative meaning of the simplex. RE are never pre-constructed units, but spontaneous in discourse. Following this, RE always need a context to be understandable (as do RS), while RD, to a certain extent, can be grasped in isolation. Finally, RE also include other-repetitions, while RD occur in continuity in a same turn.

For instance, in baby-talk, words like "mama" or "papa" are dealt as reduplication if they are considered as conventionalized lexemes composed of two similar constituents (ma/ma), but as RE when they are pronounced by the growing baby who wants either to imitate his parents (say "*papa*"- Papa) or, let's say, to train his pronunciation by consistently saying *Pa Pa Pa*. If reduplicated words are indeed lexicalized, it does not mean however that they are fundamentally of different nature than RE. Both have effects on a pragmatical level (see chapter 4).

2.3.1 Functional similarities between RS and RE

I have inserted below a synthesis including the main functions of RE. The table is drawn from Bazzanella (2011) who studied the connection between redundancy, repetition, and intensity in discourse. According to him, RE "besides being a useful cognitive device (as a simplifying/clarifying device, filler, and support both for understanding and memorizing), an efficient text-building mechanism, and a widespread literacy and rhetorical device, is a powerful conversational and interactional resource" (p. 249). RE is indeed a useful device in conversation:

Macro- and micro-functions of repetition.

1. Cognitive
   1.1 R facilitates language production
      1.1a R facilitates comprehension
      1.1b R marks comprehension
   1.2 R functions as memory support
   1.3 R draws attention to a specific item
   1.4 R is used to check knowledge
   1.5 R is used to plan one's talk
2. Textual
   2.1 R establishes cohesion between adjacent utterances
   2.2 R refers to shared information which can be used as topic
   2.3 R is used to answer (routinely)
   2.4 R is used to imitate
   2.5 R is used to comment upon (even in an aside)
   2.6 R is used to resume and paragraph
   2.7 R is used to restate one's position
   2.8 R is used to quote
   2.9 R is used to stress a point
3. Stylistic
   3.1 R stresses parallelism and correlations
   3.2 R emotionally amplifies and marks involvement
   3.3 R increases emphasis
   3.4 R marks the duration
4. Argumentative
   4.1 R increases argumentative force
   4.2 R is used to claim and counterclaim
   4.3 R is used to increase specificity
5. Conversational
   5.1 R makes the conversation advance easily
   5.2 R is used to take the turn (also to take the turn again after an interruption or an overlap)
   5.3 R is used to keep one's turn in spite of interruptions
   5.4 R is used to hold the floor, as a filler
   5.5 R is used as an attention-getting device
   5.6 R is used as a floor-yielding device
   5.7 R is used in order to inform a newly-arrived interlocutor
   5.8 R is used in seeking acknowledgement
6. Interactional
   6.1 R marks agreement
      6.1.1 complete agreement; R ratifies the preceding turn/recognizes correctness, demonstrates appreciation, enjoyment
      6.1.2 partial agreement
      6.1.3 R is used in order to show that one shares attitudes, opinions, etc.
      6.1.4 a suggested lexical item is repeated in order to adopt it in the speaker's own discourse
   6.2 R marks attention, involvement, and participation in the interaction
   6.3 R is used as 'information-receipt' (also to self-inform)
   6.4 R marks surprise
   6.5 R is used as a request for clarification
   6.6 R is used as a request for confirmation
   6.7 R is used to 'reverse' the direction of the question or of the utterance
   6.8 R is used as a request for correction of the entire utterance or part of it
   6.9 R marks disagreement
      6.9.1 R displays ironic or humorous aims
      6.9.2 partial disagreement
      6.9.3 disagreement or opposition
7. Ethnic
   7.1 R is used to stress group identity
   7.2 R is used in performing rituals

Bazzanella 2011, p.252

In the same manner, there are countless ways of reporting, each one aiming at a special result. Those effects sometimes overlap with one of these "macro- and micro-functions" of RE. I enumerate below several features found in articles about effects of RS and put them each time succinctly in connection with their equivalents (found in the table) in RE. The following comparison is not exhaustive and only aims at showing the consistent link that relates both phenomena. Functional equivalences are cues inciting to research what they concretely do in the dynamics of conversation, i.e. on which pragmatical area they are effective.

A) Substantiating and authenticating an assessment

Non-narrative quotations are used to substantiate and authenticate assessments. They do this by introducing a congruent evaluation of the same assessable on a prior occasion which is reenacted to demonstrate the here-and-now assessment and is to be taken as supporting and/or strengthening it. (Couper-Kuhlen 2007, p.100)

Here is an example (*ibid.*, p.99) of Emma and Lottie's phone call at Newport Beach (see Appendix):

(9)

```
1 Lot:        … afte[rnoon] yihknow when ih when it isn'so ho:t.
2 Emm:             [ Yea:]h.
3 Emm:        YE : : :s I would tu- It's yis been beautiful though oh
4             [it et's:] Bud] she  its so smoggy uptown it's js terrible=
5 Lot:        [Oh hadn'] it? ( )
6 Emm:        ='n Do:n came down this morning's siz it's jis terrible'e
7             siz God it's beautiful down here
8 Lot         Ye::ah.
```

Emma confronts Lottie by saying that, despite the heat, the weather is beautiful. To support her assessment, she first quotes Bud, who has claimed that it's smoggy uptown. Indirectly, his words support Emma's assessment that the weather is (after all) beautiful at the beach, as contrasted to the town. Next, Don, who had the chance to have been in both locations successively, is quoted to upgrade the assessment that it is far more beautiful at the beach than in town.

Point 4.1 in the table is close to this function. Any RE can be argumentative, because, as previously said, they steer the inference, as in e.g. "he is *dead* dead". The RE as an inherent intensive and convincing effect (I'm not joking: he's really dead, not only badly hurt), or in "Your father writes about *terrible* terrible things, and I don't want you to know anything about that." (Believe me: he writes terrible things).

B) Hedging a dispreferred action:

Verbs of thought are especially used to account for dispreffered/disaffiliative action:

> Quoting one's own thoughts […] can be a means to warrant some accountable action, or the report of some accountable action or lack of it. The former calls for justification because of its dispreferred or disafilliative potential, the latter because negative reports of not doing something amount to admissions and thus necessitates defences. […] Quoted thoughts can, and often do, incorporate expressive displays of affect, hallmarking the inner state which accompanies or accompanied the taking of a decision (*ibid.*, p.110).

An example would be: "Mr President, please excuse me, I may be mistaken *but I thought that* when votes were about people they had to remain secret, which has not been the case."[11]. The disafilliative assessment is embedded in a reported thought, which gives to the speaker the right to contradict his hierarchical superior. Verbs of thought have the advantage to exhibit affects, states of mind or intentions, which would otherwise remain unsaid. This is precisely what RE strives to do, e.g. "But I wasn't *sure* sure, you know, so I didn't do it." or in French "j'allais pas super super (ou: trop trop) bien, donc je ne suis pas venu." Erasing the RE would have the same consequence as removing the verb of thought, and would make the statement less admissible. Someone who is seeking the adhesion of the interlocutor often draws on RS/RE.

"Hedging" corresponds to point 5.8 in the table, but stays very close to "impropriety" (see below), as both make strategic use of quotatives to seek affiliation.

C) Joke session

Enacted sequence of utterances create a high degree of affiliation (Holt 2007, p.73). The original topic of discussion is held back in favor of a second, playful one with the help of a framing message (Hopper & Glenn 1994), a *key* in the terms of Goffman, which is a metacommunicative signal

> by which a given activity, one already meaningful in terms of some primary framework, is transformed into something patterned on this activity but seen by the participants to be something quite else. (Goffman 1974, p.43)

A RS enacted between participants is an example of "brackets" that transform the original interactive activity into a joking "drama" (see fragment 1). As we have seen in chapter one, a reported utterance never extends beyond one TCU. There is therefore no such thing as an "enacted scene", rather a sequence of enacted RS embedded in a hypothetical, playful scenario. "Play is reaccomplished in every utterance" (*ibid.*, p.32), i.e. play is the product of an emerging interpretation whereby two partners co-construct a situation by building on each other's utterance.

RE can also signal that a conversation has been reframed in another *key*, sometimes a mocking or joking one (Norrick 1994, p.18-19):

(10)

Lee: It's a good tape recorder. It's a nice one, hunh?
Nan: It's a beauty
Lee: Beauty. Yes. Top of the line. State of the art.
Nan: Huh huh/huh.
Lee: And the cutting edge
Nan: All of the above
Lee: All of the above. Oh::. I lo::ve it. Can I ha::ve it huh huh heh heh he he/he
Nan: Huhhuhhuhhuh
Lee: It's beautiful, oh my God.

/ = indicates a point at which next speaker interrupts

This fragment includes many RE, which "suggest that both speakers have adopted a mocking key and have consciously expropriated voices not their own and at odds with the actual interaction. The discrepancy between the speaker's own beliefs and their verbal performance creates the script clash characteristic of jokes and joking" (p.19). RE here fulfill the exact same role as enacted RS in multi-turn sequences (which are very often humorous). A comical or ironic situation is conjointly constructed in both cases. Joking situations, alongside parody and irony, match the point 6.9.1 in the table ("RE displays ironic or humorous aims").

D) Impropriety

Speaking more overtly or with a register inadequate to the speech communication without shocking the audience often requests a RS. Framing a speech with a quotative suspends namely imposed language code and social restrictions. "One can breach the rules of appropriate speech behavior provided such unexpected behavior is placed within a rule-satisfying frame." (Maynard 1996, p.214). The intimacy provoked by the emotional-affective dramatizing effects of DRS is enabled by the distancing effect created by the mulivoicedness inherent to RS. An example (in Japanese) would be:

> Mattaku hito o bakanisuruna   yo tte yuu n   da.
> really   person O make-fun-of-NEG IP QT say NOM BE
> 'I say (to you) – really, don't make fun of me.'

Maynard 1996, p.215

The formal context in which the utterance is embedded does not allow participants to make such a straightforward remark. However, making it a quotation empowers the speaker to express his inner feelings without breaching the norms. RE allows too to say things more bluntly, e.g. saying to a college who comes late at a reunion: "Je ne trouve ça pas très sérieux sérieux" ("very serious") in order to show without offending him that his delay has not been well appreciated.

E) Parodying

Someone quoting himself can "nullify the effectiveness of his own speech act" (Maynard 1996, p.221), or weaken the effect of his words. Referring to one's own speech and presenting them as caricatures is a way to minimize them in order to counter an awkward situation, e.g. a possible offense towards somebody or an institution, as shown in the example below:

> (B comments on the geographical features of the university)
> B: Uun/ daigaku   tte yappari kono kurai
>   uhh university QT after all this   much
>   hirokunaku cha/ ⌉ LAUGH ne to   ka itte./ ⌉
>   must be large   IP QT Q   say
>   (A: ⌊ LAUGH)   A: ⌊ Un)
>   (Yeah.)
> 'Well, universities should after all be as large as this, (LAUGH), you know,
> (Lit. saying something like) just kidding.'
> (casual conversation)

Maynard 1996, p.221

B realizes that she is praising another university than the one to which she belongs. Knowing that she is recorded, she prefers to get away with parodying her own words in order to mitigate their

impact. Parody is also often used in RE, in bickering or other situations of disagreement, when somebody repeats with an annoying voice what the other just said, in order to ridicule or trivialize his words. On the contrary to "joking", parodying happens much more in situations of disagreement than of affiliation. This is not as surprising, as in situations of conflict, each one wants to impose his point of view to the other. The difference is that it is made by defeating the belief of the other first, i.e. to convince him that his opinion is ridiculous.

F) Epistemic authority and subordination in assessing

According to Clift (2007), a "speaker's differential rights to assess referents are tacitly encoded in who produces an assessment first, and who second" (p.128). These rights can be revoked by several devices, e.g. tag questions (isn't it?), evidential weakening (it seems, it sounds), negative interrogative (isn't it beautiful), or, on the contrary, asserted: reporting a past speech event, for example, provides "a powerful evidential display of having reached that assessment first." (Couper-Kuhlen 2007, p.128) In the extract (2), Marsha seeks to defeat the implication that Tony has the first right to assess by using "epistemic upgrading" indexers, i.e. devices used to claim epistemic priority above someone else. Three of them are present in the extract: (i) *oh*-prefacing, (ii) increasing Tony's word "mad" in "disgusting", and finally (iii) the reported utterance, introduced by "as a matter of fact". A RS is the best resort to providing evidence for what is assessed.

RE is also a way to prove one's authority: by doing so, somebody signals that he will stand by his position. Adding an intensive *very* when repeating somebody's words is also a marker of epistemic authority, for instance, the typical Dupond and Dupont quote: "It's beautiful today/ I would say more: it's *very* beautiful today". Dupont, who systematically duplicates Dupond, tries consistently to intensify assessments he never reaches first. This happens more generally when a RE is upgraded: "This is beautiful! - Yes, this is really beautiful", or even in adjacency pairs which do not display any clear assessment: "Do you like it?- I like it very much". Each time, the respondent wants, as it were, to take the lead. By adding "very much", he expresses that he had already "liked it" ever before someone asked the question to him, i.e. that he had already the emotion inside of him before being prompted to assess his feelings overtly.

According to this short review, it can be surmised that RE and RS share very similar functions. At each point, a speaker orients the way an utterance has to be understood, what I call "constraining inference". As we will see in the next section, this principle must be relativized, as it does not account for many other cases of RE.

2.3.2. RE and inference constraint

I have already underscored the relationship between inferences and RS. But what about RE? Does it also manage inferences? I have prematurely hypothesized (p.17) that it is indeed the case. RE shows how the speaker relates to the repeated word and how he wants the hearer to interpret it. In other words, it steers the inference procedure. Dansieh (2011) supports also that view. He illustrates it with an example drawn from the translation of the Bible into Waalii:

'Bi-tunihi 10:38, WNT (Translation into Dagaare/Waalii)

Bila   ni   Naangmini ang   daang buo a Sipegi'yiruu   a   paahi kpiengu
How that chief-god   PAST pour   DEF   spirit-wash-clean CONJ add   power

a   ing   'Yesu ang 'yi   Nazare   ingang. Ung   da   mang yo
CONJ put-in Jesus who come-from Nazareth body   And-he PAST HAB roam

yuora a   tuna   sung, nyaara   bamahi
REDUP IMPERF do.IMPERF good making-well these-ones

jaa Sitaani ang dogiha,   'bojung Naangmini daang 'biye   u seng.
all satan be punish.IMPERF because chief-god PAST be.PERF his place

p.170

According the author, in the example above, the main verb *yo* is repeated with the infinitive joined to it. In this context, the original meaning of *yo* "roam", develops into "roam about". This meaning is not encoded in the denotation of the word itself, but is contextually inferred: the preceding habitual tense marker *mang* supports this interpretation because it denotes repetition and continuity. The author concludes: "By doubling the verb, the translator gives some explicit indication of the way the utterance is to be processed, that is, the direction in which relevance is to be sought." (p.170). RE displays a range of possible interpretations, only one of them being selected thanks to several parameters (e.g. cotextual elements) that will lead the addressee to guess the optimal relevant interpretation.

Rossi and al. (2014) have pointed out what kind of inferences are made when people are confronted with instances of Creative Total Reduplication (what I name RE, e.g. It's a *little little* cat). They confirm that CTR "are not determined from a linguistic viewpoint", that they "can be interpreted in different ways, depending on many factors, notably the context", and finally that they "may be additionally used to convey an affective inference, namely an inference about an affective meaning" (p. 353). "Affective" RE are often found in oral narratives, as a way to involve the hearers in the story told. Let's take some RE found randomly in books written in French:

(1) Sur quoi le démon, s'élançant après lui, *agrippa agrippa* les mains à la croupe du cheval et la voulut retenir, mais parce qu'il y avait été appliqué de l'huile, il la laissa *glisser glisser* et ne put la saisir.

38

(Histoires qui sont maintenant du passé, 1968, trad. Bernard Frank, p.177*)*

(2) Les murs étaient peints en blanc, *lisses, lisses*, et brillants. Dans cette immensité, il y avait une aiguille, *fine, fine,* dure, brillante sous la lumière. Cette aiguille dans ce vide me faisait un peur atroce.

(M.A. Sechehaye, *Journal d'une schizophrène*, 11th edition, 2003, p.13)

Upon reading the RE, one visualizes the teller and can easily imagine which gestures he would have performed and which emotions he would have ostensibly expressed while repeating. This is not that surprising, as the first extract is drawn from a tale that was once orally transmitted, what implies that the story was told in real live. The second extract is the recount of a dream, what suggests that the author described something that she personally experienced in her sleep. In recent years has been developed the so-called *Empathy theory of dreaming*, which proposes that "the sharing of dreams has an emphatic effect on the dreamer and on significant others who hear and engage with the telling of the dream" (Blagrove and al. 2019, p.1). The two RE in (2) could suggest that a recount of a dream, at least in the case above, is an enactment, what is further supported by the fact that a speech enactment often aims at sharing a point of view and triggering empathy (cf. the *like+* "…" construction is used when one wants that the partner empathically conform to one's point of view). Until now, I have associated speech enactments exclusively with the principle of inference constraint. Do these "narrative" RE entail some kind of inference? They suggest indeed that suprasegmental features, such as intonation, has to be "inferred" by the reader in order for him to understand the affective/expressive value that it conveys. This active recovery of the original intonation is exactly the same as when one reads a reported speech. Both of them have to be integrated in a real-time speech act to be effectively understood. Apparently, there is also a second-level: in the experiment of Danni et al., which "consisted in a delayed verification task involving a judgment of consistency" (p.356), test sentences were orally submitted to the participants:

> Results have strongly confirmed the first hypothesis: on the one hand, participants having heard a CTR made significantly more inferences than participants having heard the simplex; on the other hand, participants having heard a CTR judged as consistent significantly more sentences containing an affective inference than subjects having heard the simplex. (p.364)

They further add: "the presence of a specific linguistic structure may indicate the presence of an additional affective meaning and generate the inference of that meaning" (p.352). The emotional impact of CTR thus increases interpretive effort.

In (1) and (2), it is indeed quite obvious that RE has an involving effect, but still, we cannot speak about inference *constraint*, above all because they are embedded in narratives and depict things beyond the concerns of direct meaning negotiation. Other exceptions are easy to find. I was told the other day by someone who wanted to share his surprise over someone else's behavior: "Ce

39

n'est quand même pas normal qu'il ne cuisine jamais jamais[12]". Is there a preferred and a less-preferred inference? A "never-*never*" and a "never-*sometimes*"? I had rather the impression that by repeating, the speaker did not want to assert anything (cf. point F., this chapter), but wanted to establish a ground for discussion (or rather for gossip). Inference constraint is therefore not the only explanation to explain all enacted RP, but it certainly participates in the negotiation of meaning.

To conclude, I have observed in this chapter some accounts made by various researchers about RS and RE, emphasized recurring aspects such as the joking and parodying, as well as the argumentative (assessing, substantiating) ones. Their various pragmatic roles show that they present important strategies in meaning negotiation. Furthermore, although inference constraints seem to play a major role in both phenomena, they cannot explain a significant number of other cases.

All these miscellaneous accounts should be understood more globally, in a consistent frame that deal with them as *speech enactments*. How does RP make all these pragmatic functions possible? How does it enable the negotiation of expressive meaning? The dynamics of interpretation goes actually far beyond inference constraint, and RE are often a cue that an inconspicuous underlying activity of perspective alignment pervades the dialogue. This less visible mechanism should be brought to light thanks to its most visible manifestation: echoes.

3. Echoing

I would like to split the general term of "RE" into standard RE, which animate the speech and constrain inferences, and *echoes*. Why exactly "echo"? Whether we repeat our words or the words of others in a specific context, we often make resonate the words inside of us. Mikhaïl Bakhtin (1981) pointed out that polyphony – or multivoicedness – pervades any dialogue. Very often, we enact through a full or partial repetition fragments of other's speech in order to co-authorize something we did not say: "Each utterance is filled with echoes and reverberations of other utterances to which it is related by the communality of the sphere of speech communication. (Bakhtin 1986, p.91). These echoes are kinds of short quotes where the speech of others (or even one's own speech, see reflexive echo 3.2) is enacted rather than vaguely "animated": it is indeed difficult to perceive echoes as "animations". Therefore, the more specific concept of "enactment" seems to fit better to account for them. I argue that all echoes are enacted on the basis that they participate in the active negotiation of meaning, (re)present an utterance, are situated in a "perspectivized" there and then, and prompt an emerging interpretive process.

---

12 "jamais jamais" is a kind of diagrammatic RE (chapter 4), and not a RE constraining inference.

Four kinds of echo will come under scrutiny: simple, reflexive, and recursive echoes. The fourth one (impersonal echoes), is particular in that it includes some RD and lies somewhere between RS and general RE.

3.1. Simple echoes

Simple echoes comprise all types of other-repetitions in the flow of the interaction. They are simple because, despite the entanglement of voices, the original author of the echo is clearly identifiable (there is no reflexivity or recursion at play, see next sections).

Fragments of an other's utterance can be repeated, in order i.a. to confirm, clarify, correct or question it. An echo "[conveys] the speaker's attitude to an attributed utterance or thought" (Wilson 2000, p.248), e.g. if John says to Mary: "Anthony has horrible tastes", Mary may assess what he said by saying:

A. Horrible tastes. [affirmative]

B. Horrible tastes! [surprised]

C. Horrible taste… [doubtful]

According to relevance theory, irony, denials and echo questions are also cases of echoic constructions. Denying implies a rejection of an attributed assessment before substituting a righter account for it. Examples are from Wilson (2011, p.22):

(1)   a. *Peter*: Oh, you're in a miserable foul mood tonight.
        b. *Mary*: I'm not in a miserable foul mood; I'm a little tired and would like to be left alone.
(2)   Around here we don't eat tom[eiDouz] and we don't get stressed out. We eat tom[a:touz] and we get a little tense now and then.
(3)   Mozart's sonatas weren't for violin and piano, they were for piano and violin.
(4)   I didn't manage to trap two mongeese: I managed to trap two mongooses.

Each time, the speaker has to deny the wrong assessment, before moving on to the correct one. The reported utterance can be overtly Marked, as in (1), or implicit, as in (2-4). This is somehow parallel to inference constraint (i.e. they are contrastive), except that there is here nothing to infer, as every piece of information is made explicit.

Irony too may consist in an attributed utterance repeated as an echo, and augmented with a marker signaling a dissociative attitude (a light modification, or an exaggerated intonation), as in the following mocking inversion (ibid., p. 21).

(5)   A. Our friends are always there when they need us.
        B. Our friends are always there when we need them.

It can already be noted that, whereas irony in hidden quotations (cf. p.17) implies that one reflexively takes perspective on one's own speech, in simple echoes the person who is ironic enacts and modifies somebody else's speech. This is potentially offending, as the voice of the first speaker is stolen and "hacked" by the second one, who distorts the original meaning and reflects it to the original producer (cf. also recursive echoes).

Following the theory of mind at the basis of relevance theory, echoic questions are also made on the basis of inferences relating to the presupposed intentions of the other. Attributed thought can therefore also triggers an echo question:

(6)   a. *Mary* (seeing Peter walk towards the door): Just a minute. You're going shopping?
b.?*Mary* (seeing Peter walk towards the door): Just a minute. Henry VIII had six wives?
c. *Mary* (seeing Peter walk towards the door): Just a minute. Did Henry VIII have six wives?
(Ibid.,p.23)

In conclusion, echo utterances have a "questioning, transforming and distancing" effect (Di Paolo, p.190). Contrary to RS, they are ostensibly different from what they repeat: ironic echoes modify usually an element, while denials reformulate the element repeated to make it match to the reality.

From all the miscellaneous cases encountered so far, it is clear that the term of "echo" is generic and refers to a fuzzy set of other-repetitions. There is however still one kind of simple echo that is needed to get a full picture of how simple echoes can be instantiated: "back-channeling" simple echoes. I understand "back-channel" in the general meaning of "a sound or sign that someone makes to show that he or she is listening to the person who is talking"[13]. Back-channeling simple echoes fulfill many roles that support and maintain like-mindedness between interlocutors: agreement, sympathy, adoption, questioning, filling of space, enjoyment, etc. "The repeaters frequently adopt words and phrases through repetition in order to better relate to the topic and maintain coherence within the talk" (Machi, International spring forum 2011). Here a few examples drawn from a corpus of conversations in Japanese (the untranslatable japanese particle ね "ne" in 2. "marks the participants' intention to identify with the knowledge, judgment, and feelings of others" (Machi, 17th sociolinguistics symposium, 2008):

1) Linking

A: I went into a **haunted house** yesterday and it was so scary.

B: Well, [I] went into the **haunted house** in Disneyland and […]

(2) Sympathy

A: That was really **dangerous**.

B: **Dangerous** (*ne*).

(3) Agreement

A: When I saw him there, I was so **surprised**.

B: (no subject) **surprised**. [14]

Of course, back-channeling in general extends far beyond RE, and includes most of the time short verbalizations, such as a single *yes* or an *uh huh*, or non-verbal messages (head nods, smiles, eye contact) indicating agreement or showing that the perspective is shared. Short answers, which can be seen as back-channels prompted by the partner, combine a *yes* with a proverb or a full RE, or just a full RE, like in the following example in Japanese, where answers to questions are always the echo of the main verb:

パン　を　食べます　か。　- Would you like to eat some bread?
pa n　o　ta be ma su　ka.　*(bread/[object]/eat/ [question])*
(pan o tahbémahs' kah)
食べ　ます。③④　- Yes
ta be　ma su.　*(eat)*
(tahbémahs')

(Garnier & Toshiko 2007, p.9)

Back-channeling simple echoes enter into a broader category of resources which "highlight the interpretive processes of the current producer" (Di Paolo & al. 2018, p.188). In a dialogue, a listener can chime in and influence the way a producer interprets his own speech. Simple echoes have the advantage to thematize the specific part they comment on (by echoing the word). Sometimes, they are even necessary, as in this nonnative/native speaker conversation:

(11)

17: Jose: no. writing is- is no, because I don't have time, time.
18: time.
19: the time is very short. shorter? shorter?
20: is uh: sho-is-is short.
(Knox 1994, p.203)

---

14 https://www.meertens.knaw.nl/ss17/contributions/abstract.php?paperID=724 (accessed on January 24, 2020).

In 20, the native corrects the word Jose is trying to utter. On the contrary, in 18, he shows him with the help of an echo that he understands what he means and that he can go on without spending more time on *time*. This extract is interesting, because it makes clear how a simple echo is used by a cooperative interlocutor to assert the speaker in his speech. In 17., the latter makes a reflexive echo (see next section), in search for self-support: by repeating *time,* he tries to take perspective on the problematic word and to keep the floor. The native decides all the same to assist him by a simple echo. In 20., however, he has to answer to a question of the nonnative: his answer is not a simple echo (a single *short*) but a hesitating *short* followed by a more assertive one. Doubling in a same turn is reflexive (see also next chapter). The native is here a bit surprised to be asked to take the floor, and repeats *short* to regain quickly self-assurance in order to answer promptly.

Traditionally, back-channels have been seen "as a means to avoid taking the floor from the current speaker", and the responding "addressees are seen as passive recipients of information, with backchannels being used to display addressees' acceptance of speakers' planned multi-turn utterances" (Tolins & Fox tree 2014, p.153). This view has been been debunked by the *proactive backchanneling theory*, which claims that listeners actively collaborate in the unfolding of a speaker's talk (*ibid.*) Each talk (especially narratives) is a joint action rather than a unilateral contribution. Backchannels in particular steer the development of the conversation itself. Here an illustration drawn from Tolins & Fox Tree (p.157), in common Jeffersonian transcript:

(12)

26 S5: Didn't Miss Lewis ever tell you about like her nephew or something
27 S6: ([no she did]) I probably forgot
28 S5: In the navy
29 S6: Mm[mm I dunno
30 S5:   [ok- ok she had this nephew that was like- he was in the navy
31   and you have to be short cause to fit in the submarine you know
32 S6: Uh huh
33 S5: like cause they only make it like a certain height and he was like
34   only 5'6'' or 8''
35 S6: Uh huh
36 S5: and then like he had a growth spurt while he was in the navy
37 S6: Uh huh
38 S5: and this is like bef- when he was twenty or twenty one and he
39   turned to like six something.

The authors comment by saying that

the two interlocutors can be seen as creating the discourse together. At each point in which new information is presented, the addressee accepted this information and displayed understanding through her use of backchannel communication. We argue that it was not necessarily S5's goal to construct a multi-turn utterance when she began her tale at line 30; rather, it was the joint process of presenting and

accepting discourse events and relevant information, through generic backchannels, that lead to the construction of the speaker's narrative. (p.157)

The author's view is very much in line with what I called "emerging interpretation". Bach-channels in general, and not only back-channeling simple echoes, display the continuous interpretative effort engaging two partners, who construct the speech together (the listener steering the conversation with the help of echoes, nodding and so on). What further strengthens this assumption is that in the fragment above, the speaker himself bolsters his words with so-called "fillers" ("you know", "like"). "You know" is revealing, because it signifies that co-participation is needed. Every speaker needs (to varying degrees) a supportive attitude from his interlocutor to keep talking. The listener, who punctuates by his reactions (expressive gestures or verbalizations) the ongoing speech, sustains him. He is a kind of mirror through which a speaker continuously produces and interprets his own speech. This retroactive process is characteristic of speech enactment. Every turn is an interactional achievement: there is actually no "turn-taking", in the sense of alternating between speaker and hearer, as each one permanently looks for the perspective of the other (what Maynard called in the introduction "perspective of becoming", see introduction), viz. looks for "becoming" the other in order to adopt his point of view.

By echoing, the listener aligns to or disrupts the way the speaker positions himself towards what he says. In a dialogue, the latter consequently does not express so much his point of view as the point of view of the listener, in a distributed way. This is why, as an example, ironical simple echoes are frustrating: the one whose words are sarcastically repeated cannot do but integrate the perspective of the other into the interpretation he makes of his speech, what can destabilize him. Consider the situation when somebody, instead of collaboratively back-channeling, draws the speaker's attention on a malapropism by imitating it or pointing it out, e.g. the frequent confusion between DRS and IRS in French. A. "Il m'a demandé *qu'est-ce qu'il* y a et j'ai…", B. [critical tone] "Il m'a demandé *quoi*?". Such interference may unsettle the speaker, because not only cannot he gain a positive feed-back that will assert him, but he also faces a critical response that he cannot help but include into the view he has of his own speech. This can jeopardize the smooth running of the conversation (the speaker stutters; coherence and cohesiveness drastically decrease).

In conclusion, simple echoes are not just reactions to what is said, but a constant feed-back that directly influences the way a speaker takes a stand on his speech. By contrast, non echoing back-channels have a limited role, because they can only display the listener's (dis)agreement or confusion (e.g. frowning, nodding,…)

45

It also happens that one interprets oneself without the support of the interlocutor. The enactment becomes reflexive. For instance, using "you know" at the beginning of a sentence without expecting the other to back-channel, or uttering "I mean" (much too) often. Echoes also have this reflexive capacity. In the next section, I will deal with such cases.

3.2. Reflexive echoes

3.2.1. Reflexive echoes: quotatives, particles, length variation

One type of RE, out of the three that I have proposed to investigate, may immediately benefit from what I have discussed so far about new and traditional quotatives (point 2.2). In reflexive echoes, new quotative *like* can be often inserted between the two elements repeated.

Let's observe first how a contrastive focus RE like *now now* interacts with *like*:

1) She wants to see me *now now*?
2) You mean, *like, now-now*?
3) I think we should go, *like now, like* right *now*.
4) I just want to settle down and have a gaggle of rug rats, *like, right now*.
(Source: reverso.net)

In the first example, *now* is merely repeated. But *like* can also be inserted before the second *now*, as in (2), or before each *now* in (3). *Right+now* in (4) amounts to say *now-now*, because both of them suggest an immediate *now*, by contrast with a delayed *now*. *Right now* and *now-now* are, as a matter of fact, interchangeable. *Mean* often appears in such context too:

(5) I mean, I need you here bad.

(6) I mean, that's *like really* taboo.

(Source: reverso.net)

The fifth example, that I draw from an English-speaking television series, does not include any RE. What's interesting is the French subtitle corresponding to it: "J'ai genre, vraiment vraiment besoin de toi". *Vraiment* would not be repeated in the translation if the original sentence was *I need you here bad* without *I mean*. Furthermore, the filler *like* is always virtually present in this type of RE: *now now* can be paraphrased *"now", like: now?*. The two components are bound by an integration relation. It means that *now now?* could be reformulated as: Does *now* belong to what we commonly refer to as "now"? Does it refer to the potential (viz. prototypical) meaning of "now"?

What can be observed from the examples above is that one speaker anticipates the partner's expressive meaning. The negotiation of meaning is partly undertaken by one person who tries to

regulate the situation by himself, either by anticipating the answer (you mean: now-*now?*), or reformulating his own words when he judges that it is not clear (I mean *X*). Momentarily, one partner takes everything in charge by adopting a split attitude. Let's take a random utterance in French: "J'étais dégouté". If I want to share this feeling, I could say: "J'étais genre: dégouté". In that case, a back-channel from the partner is essential before I resume my speech (e.g. *oh really?*) But I can use a reflexive echo (in 1.) or particles (2 and 3) so that I do not need this back-channel. French displays a great array of particles in addition to the original reflexive echo:

    (1) J'étais dégoûté, genre **dégoûté**. (the typical "echo")

    (2) J'étais (genre) **mais** dégoûté. (a particle)

    (3) J'étais (genre) dégoûté, **quoi.** (another particle)

    (4) J'étais dégôuté, **mais** genre **dégoûté quoi**. (the addition of them)

Each particle is pretty similar to Eng. y*ou know*. By saying *you know* and *like* in the same sentence, one simultaneously initiates an enacted speech, but reflexively regulates the joint meaning negotiation by oneself: "I was like disgusted, you know". The partner is presupposed to know what he means, so that he does not have a say in the matter, or is prompted to agree without being fussy. In other words: in "I was like disgusted, you know" the speaker actually does not need the appreciation or the point of view of the other. Sometimes, the quotative itself or suprasegmental features signal the reflexive function, in addition to the echo itself:

 (13) Corpus : **Bielefeld _ situations de contact – cadre universitaire**

```
G   (../..) a bien sûr euh juste avant (0.6) euh: m- (0.3) ce:: (0.3)
    pourquoi rendre la chambre (0.8) qui est dit de façon euh: (0.9)
    disons désa- désagréable
D   hm hm
(0.5)
G   e:t (.) j` m` demande si: (0.4) cette réaction: (2.5) hm:- désagréable
    (1.1) n'aide pas un peu doris (0.7) à formuler ses choses de [façon]
    trè:s très précise [((rire))]
K                      [hm hm]
```

In the first turn, G uses the reflexive echo augmented with a quotative "disons désa-désagréable". "Disons" subsume all participants, it implies that the speaker regulates the meaning by herself (not **I'd say*, but *let's say*). She is uncertain about how to encode her ideas, but does not want to engage the interlocutor in meaning negotiation, so she chooses to engage in the process alone: "Let's all say 'unpleasant'/ 'let's all agree on it'. G tries to compensate her uncertainty by resorting to a *disons,* but it is not enough: she stumbles directly after it on the onset of *désagréable,*

in order to repeat it more assertively. The quotative and the RE are here to transform the stuttering speech into a self-performed echoic enactment.

Extending the first element of the echo is also reflexive. In "très:s très", the first "très" is a signal that the speaker walks on eggshells and that meaning must be urgently co-constructed, but the echo completes the process without that it even begins. The accompanying laugh shows that G is embarrassed after having chosen to close the potential negotiation that she himself had made possible. At that point, a minimal feed-back is requested from K, who validates with a simple "hm hm", showing that it is all right, i.e. that no collaborative principle has been broken. Another example of length extension:

(14) La caution
DIR   vous faites de la comptabilité vous aussi/
STE   non::: (0.2) non [pas du tout]
DIR   [et ben vous] faîtes pa`ce que c'est vraiment:
STE   ((rire))
DIR   vraiment terrible hein

"Non non" and "vraiment vraiment" have a first, extended simplex, while the reduplicant is short. Both cases are similar to 'très:s très". This time, this is not the speaker who laughs, but the listener: DIR initiates an enacted speech by extending the first "vraiment", and STE, by his laugh, allows DIR to take charge of it reflexively, and not conjointly with his interlocutor, who presumably does not want to co-enact the speech. This last assumption is substantiated by the fact that STE uses the same strategy: the first "no" opens a potential negotiation, but the second one closes it directly after. It is therefore quite fair that he lets DIR make the exact same thing. Note also the "hein" ("eh") of DIR, who asks for a validation. A reflexive echo is like a document that one writes in the name of somebody else (with his agreement or not, cf. the "let's agree"), only to let him sign (validate) at the end.

3.2.2. Adjacent reflexive echoes

Quite paradoxically, a reflexive echo can be be produced by two people at the same time. To understand it, we should think about the favorite expression of Dupond and Dupont in *Tintin*: "It's beautiful today/ I would say more: it's really beautiful today". Dupond makes a reflexive echo with the support of a like-minded twin. Both are individuated, what does not prevent that they share a same TCU comprising a reflexive echo. Here is a real-life example of the The Dupond & Dupont effect:

15. Video game session

RAP c'est un angle fermé:/
(1.4)
RAP hi::://
LUC désoLÉ\
RAP non/ ça va\
(0.8)
RAP bien/
(0.4)
RAP devant/ on est plusieurs\
(2.4)
LUC .h::
(0.8)
LUC [ça la:gue mais c'est horrible]
RAP [mais ça lague en plus ouais]
(2.9)
RAP °c'est bon c'est bon (ça)°

RAP anticipates the words of LUC and talks simultaneously. RAP reflexively echoes "ça lague" (Eng. "it lags", "it slows down"). An adjacent reflexive echo is signaled by the upgrading *en plus* (Dupont' "I would say even more"). Another particularity is the *mais* that both of them insert in their respective utterance. Let's put them in parallel:

ça la:gue | mais | c'est horrible
          | mais | ça lague en plus ouais

The reflexive echo seems to have a similar function to the intensifier. LUC upgrades the assessment with the reinforcement: "c'est horrible", while RAP upgrades it with: "en plus". The core statement "ça lague" is taken over by each one of the two teenagers, i.e. they individually "perform" the exact same turn. In the Dupond & Dupont expression, Dupont also asserts himself with the help of a quotative evidencing that he co-constructs the utterance ("I would even say:"). The quotative enables the second speaker to connect the turn to his own linguistic agency ("*I* speak").

One may also make an adjacent reflexive echo single-handedly:

(16)

A   Bon ça t` laisse pas euh:- j` veux dire c'est pas euh: Mort à Venise
Hein/ (.) c'est aut` chose c'est plus commercial (.) mais tu passes un
Très bon moment:/ c'est beau: puis t` apprends des trucs/

L   hum hum

A   On te montre l'opéra baroque de l'épo:que
(.)

L   Puis ça doit êt` beau

A   C'est très beau (.) très très beau (.) i` jouent bien les mecs (1.8) non
C'est bien fait (.) ((en bâillant)) t` apprends des trucs sur l'époque/
(.)

A. plays "Dupond & Dupont" at the same time. It it is due to the absence of validation from the partner (to hark back to my metaphor, this time A. writes *and* signs the document). Such mono-adjacent reflexive echo often happens when for instance the discussion dies out, but nobody dares to put an end to it (e.g. by saying goodbye). Back-channels become rare, and one partner out of the two is given an extended turn which he tries to fill in by playing Dupond & Dupont. In the discussion, both partners are tired (A yawns) and do not really want to expand the discussion.

I suspect that this of echo is not only gradient (very very), but takes also the form of whole phrases with open slots filled by synonyms:

(17)

M   donc on peut: ici on peut di- non ben ça c'est plutôt une conclusion,
si on dit
P   non /ça ça c'est un argument pour/
M   /non mais je veux dire/ non non ça ça oui oui oui mais je ce que
j'ai pensé c'était que (0) si on:, si on dit que c'est indispensable,
si si on dit que c'est nécessaire (0) euhm:, de de faire des travaux
à la maison tout seul (0) il faut qu- en même temps euh: diminurer
les les les horaires à l'école, (../..)

M. tries to counterbalance her uncertainty by numerous reflexive echoes predicted by the "I mean", reaching their peak with the adjacent reflexive echo: *si on dit que c'est indispensable/ si on dit que c'est nécessaire*. M. tries desperately to consistently set out her ideas, but does not achieve it. Moreover, she is in conflict with P., who does not grasp what she is trying to say, and therefore cannot back-channel. M., who has a clear idea in mind but does not know how to properly express it, makes successively a reflexive echo (si on:, si on dit), during which she opens a space of negotiation that she decides to tackle alone, and an adjacent one (*si on dit que c'est indispensable/ si on dit que c'est nécessaire*), because no response is coming. Another example below:

18. Corpus : **Repas ~ conversations entre étudiants - lyon 2006**

J   tu penses que toi tu auras tes résultats d'examens en en juin
M   non on s'est dit que ils pourraient être sympas ils pourraient nous
les donner avant les vacances de pâques histoire qu'on ait une sè- une
phase de révision très agréable quoi

This time, M echoes "ils pourraient" because the first part "ils pourraient êtres sympas" is much too vague for J to validate. The second part "ils pourraient nous les donner…" aims at giving him the chance to understand precisely what M means.

To conclude, echoes provide an essential support for reflexivity, because they help speakers to adopt a "split attitude" that sustains the inner dialogism necessary to think, what Di Paolo defined as "linguistic agency":

> Linguistic agency has the power of thematizing the body, and for this reason, of displacing itself from it. The body is acted on as if it belonged to another agent, one we engage dialogically, and for this reason, implicit in this form of engagement is a role separation between a more active, leading facet and a more passive, led facet of the linguistic body. (Di Paolo, p.196)

Linguistic agency is reached once one knows that the linguistic knowledge that has been gained can be interpreted in a community, that one can make oneself understood with it. But beyond conventionalized patterns, meaning is sometimes endangered. Through reflexive echoes, a virtual reflexive interaction is recreated, through which a speaker convinces himself that what he says can be interpreted. In simple echoes, the partner is necessary to do that; in reflexive echoes he is reduced to a validating role. But what if now, a speaker does not integrate the voice of the other or the self-as-an-other, in other words that he enacts it *without* embracing it? This is where the concept of recursivity comes into play.

3.3 Recursive echoes

This kind of echo can be easily found in the literature. Let us first observe two examples drawn from novels:

(A)

Ich sah ihm stumm zu.
„Rauchen Sie?" fragte er und, als ich den Kopf schüttelte: „Darf ich Ihnen einen Kaffee… eine Erfrischung…? Vielleicht einen Cognac?"
„Nein, vielen Dank!" erwiderte ich. „Ich möchte nichts, Herr Professor! Nichts außer Ihrer Stellungnahme zum Fall Westphal!"
„**Fall! Fall!** Wie das klingt!" rief er und mimte den Empörten. (Martin, *Der Rest ist Sterben*, H p. 56) (Larrory-Wunder 2016, p.83)

(I watched him in silence."Do you smoke?" he asked, and, as I shook the head: "May I offer you a coffee, or a drink" Maybe a cognac?No, thanks!" I replied. "I don't want anything, Professor! Nothing but you opinion about the case Westphal! -**Case! Case!** How funny it sounds! He shouted while pretending to be upset. (my own translation)

(B)

Il y a longtemps que vous vous passionnez pour la nature? Demanda Brancardier (…)
- La nature, la nature, c'est vite dit
(Watine 2015, p.58).

In literature study, this type of echo is called "reprise écho-dissensuelle" (Watine 2015). When repeating, one can distance oneself from the words of another by sending them back to the original author. As in reflexive echoes, quotatives or metalinguistic signals are often juxtaposed to the echo (Fr. C'est vite dit, Ger. Wie das klingt, Eng. It's easy to say,…), in order to put the repeated utterance in perspective (i.e. to detach oneself from it, to keep them at a distance). I call them *recursive* because the one who echoes integrates the voice of his interlocutor without mixing it into his own. I would like to illustrate this by citing an extract from a Japanese folktale which features a rare case of recursive RS. Recursive echoes could be considered as a contracted type of such a recursive RS:

> Or, en ce temps-là, il y avait un homme qu'on appelait le ministre du Hon-in. Dans sa maison, il y avait une femme Madame Jijû. C'était une personne en service de cour, d'une figure et d'une façon d'être admirables, aux dispositions d'esprit plaisantes. Comme Heichû fréquentait ordinairement chez ce ministre du Hon-in, il entendit parler dela façon d'être admirable de cette Jijû et, durant des années, de la façon que l'on ne saurait dire, au point qu'en échange il eût donné sa personne, il la courtisa; ce à quoi Jijû ne fit même point la réponse d'une lettre. Alors Heichû, se lamentant et se désolant, écrivit et lui envoya une lettre dans laquelle il disait: "Daignez faire voir ne fût-ce que ces deux seuls mots: 'J'ai vu'." Ainsi, à maintes reprises, au point, pourrait-on dire, qu'il en pleurait pleurait, il écrivit et lui envoya.
>
> Comme son messager rentrait porteur d'une réponse, Heichû sortit à sa rencontre en se heurtant aux objets, prît en hâte la réponse et la regarda: or voici qu'en-ce qu'il avait écrit et envoyé dans sa propre lettre: "Daignez faire voir ne fût-ce que les deux mots *J'ai vu*", on avait découpé ces deux mots *J'ai vu* et qu'on les lui envoyait, collé sur du mince papier d'écorce! Quand Heichû vit cela, le dépit et la tristesse, qu'encore plus il éprouva, furent au-delà de toute limite. (Histoires qui sont maintenant du passé, p. 197).

Heichû asks the courtesan to report the words that he has put beforehand into her mouth. He wants her, under the disguise of a simple request, to reproduce an utterance that she has not authorized, but that has been all the same arbitrarily attributed to her. Her function is then to recognize that she said it. By recursively recycling the attributed utterance of Heichû, she at the same time produces *and* reflects the "I saw" back to the original author without taking responsibility in regard to the content of the utterance. But in order to do that, she needs to change the original format of the letter and replace it with thin bark. To make it more illustrative, I drew below some concentric squares. Each square correspond to one missive:

I saw
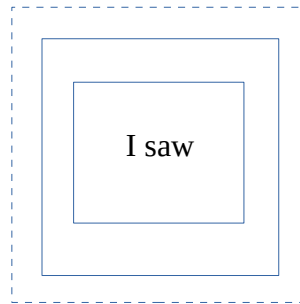
Even if it will not happen in the present case (it would be meaningless), nothing should prevent Heichû from sending back the letter on, let us say, a thicker bark. Nor the courtesan to repeat one more time the operation with a still thicker one. The process could loop endlessly. This is made possible because the reported utterance is each time reflected, from the first missive to the last. Actually, nobody says "I saw", because each one keeps passing the buck to the other: "you speak – No, YOU speak – No, you do so, etc.".

Recursive echoes act similarly. Below, a concrete example in a conversation conducted in Finnish (echo on line 9):

(13)

6. T. Mitä se sitten on?
7. A. Se on nahkaa. Vanhaa nahkaa. Viekää se pois! Heti! Ja sanokaa, mitä tuo on?
8. T. Se on kahvia.
9. A. Kahvia kahvia. Mutta millaista se on?
10. T. Hyvää kai. Tämän ravintolan kahvi on aina hyvää.
11. A. Ja kuumaa?
12. T. Ja kuumaa, aivan niin.

6. W. What is it then?
7. C. It is leather. Old leather. Take it away! At once! And tell me what that is?
8. W. It's coffee.
9. C. Coffee, of course. But what's it like?
10. W. Good, I suppose. Coffee is always good in this restaurant.
11. C. And hot?
12. W. And hot, yes.

(Aaltio 1987, p.71)

The RE of *kahvia* (coffee) indicates that A. embeds an extraneous word in his speech. He isolates *kahvia* from the waiter's utterance, copies it, pastes it into his own utterance, and sends it back: "-Se on kahvia – Kahvia kahvia.". In the translation of line (9), the effect of the recursive echo is rendered by "of course"; a reminder that English is more prone to using phrases such as "of course", "If you want", "you said it", "ok, but" than echoes. In French, it's quite common to find recursive echoes in sentences of the kind "I'm not X-X, but", for instance when Barthélémy Diaz, the major of Mermoz-Sacré Coeur, says: "Je ne suis pas malade malade, sinon je l'aurai dit »[15]. To underscore the fact that he never said that he was ill, as the media did, he repeats *malade*. He produces the word *malade,* but he is not the author of it.

---

15 https://thieydakar.net/video-barthelemie-perd-30-kg-pourquoi-je-suis-maigre/ (Accessed on April 12, 2020).

Recursive echoes could be mapped onto the "parodying" effect of RP. In the example of Maynard (p.36), the student produces and animates her speech at the same time ("I say that, but it's not me who says it"). It's easy to imagine someone uttering a recursive echo while taking overtones, in order to parody the repeated fragment, e.g. by saying "malade, malade" or "la nature, la nature" with lengthened vowels, dramatic gesturing or by making an overall satirizing imitation (in the extract in German, the professor "mimt den Empörten"; he sends back a negative portrayal of his interlocutor by repeating the rather high-sounding "Fall"). This is supposed to let the other know, for instance, how ludicrous he sounds. In every case, a paralinguistic feature is also added to the RE to show that it is a parodying performance linking production and animation. We have seen that ironic simple echoes also have a parodying effect. The difference is that in the latter case, parody is made through distortion of the original utterance, while in recursive echoes, it is made through rejection, often going with disregard.

3.4. Impersonal echoes

We have seen that personal echoes are bound to a context. They involve the participants of a conversation (the *I* and the *you*) and achieve particular pragmatic effects. Another (apart) category of echoes are impersonal echoes, strongly associated with imagic RD, including rhyming/ablaut RD, as they are typically found in infant-directed talk, onomatopoeia, ideophones, and sound symbolism in general. They considerably differ from personal echoes, because (i) they are impersonal and therefore (ii) are not bound to a deictical space, so that (iii) they can be understood in isolation. Also, they are often (iv) conventionalized lexemes which (v) are not necessarily repeated. If comparing impersonal with personal echoes would not immediately pay off, it seems that the former share more similarities with RS, starting with quotatives. In many languages, quotatives often precede (optionally or not) the iconic element, e.g. in Xhosa (Andrason 2017, p.141), where 'ukuthi' means to say, 'thi' being the reduced form:

(1) (ukuthi) cilikithi 'rush out unexpectedly, rise up suddenly'.
(2) (ukuthi) cithi 'coming out, rising suddenly into sight'.
(3) (ukuthi) cubhu 'feeling lazy, lethargic'.
(4) *cwaka*; to be silent: 'Lixesha lokuthi cwaka'. (It is time to be silent [literally: It is time to say cwaka]).
(5) *gqi*; to suddenly appear: 'Bathi gqi abelungu eAfrika'. (The white people suddenly arrived in Africa. [Literally: The white people said gqi in Africa]) .

The Japanese particle *to* also both introduce quotes (RS or RT) and sound symbolism*:*

I. Quotatives:

a. *Kare wa asu kuru to itta* He said that he will come tomorrow.
b. *Rainen nihon ni ikou to omotteiru.* I am thinking of going to Japan
next year.[16]

II. Sound symbolism

a.   dareka   ga   tama o   gorogoro to   korogasi -ta.
     somebody Nom ball Acc Mimetic Comp make-roll Past
     'Somebody rolled a heavy ball.'
   *'Somebody heavy rolled a ball.'
b.   dareka   ga   mizu o   ba   tto   mai   -ta.
     somebody Nom water Acc Mimetic Comp sprinkle Past
     'Somebody sprinkled a large amount of water.'
   *'Somebody heavy sprinkled water.'
c.   dareka   ga   mizu o   pa   tto   mai   -ta.
     somebody Nom water Acc Mimetic Comp sprinkle Past
     'Somebody sprinkled a small amount of water.'
   *'Somebody light sprinkled water.'

Kita 1993, p.403

By using an onomatopoeia, an action is traced back to a subject with an intention: *it* said something to *mean* something. What if, for example the difference between: "the dog barks", and "the dog makes: woof"? "The dog barks" is just an information that one communicates to somebody else, the sentence has just an exchange value. But if the dog makes 'woof', then the dog intents to communicate something in a particular situation: from an object, he becomes a subject (on an existential level, not a grammatical one). The same applies to non-living things: if a smoke alarm "goes off", it remains a device passively reacting to the presence of smoke (it's programmed for that), but if it makes: "riiiiiiiiiiiing" then it expresses something: from an object that passively reacts, it becomes a subject whose intentional behavior can be evaluated, for example, one could judge that it's a *stupid* device that *bothers* everybody as soon some smoke hovers in the air; it has consequently a state of mind, (because it acts by itself, and does not only "reacts"), according to which other people can judge him (it is "stupid"). Giving intentions is a basic property of images altogether, not only ideophones, for example: cave paintings depicting animals made by prehistoric people did not have aesthetic purposes, they were rather a hunting technique: by drawing them on walls, our ancestors had the sensation to explore their subjectivity. Exploring subjectivity through pictures and, therefore, guessing intentions of their preys (e.g. guessing where a herd will decide to take shelter) was fundamental and compensated largely their lack of smell (Stépanoff 2019, p.47).

---

16 Namiko Abe, "Japanese particle: to", https://www.thoughtco.com/japanese-particle-to-4077331 (updated February 25, 2020. Accessed on April 25, 2020)

More generally, children often draw pictures. The reason lies in the fact that their world is animated, i.e. they relate intersubjectively to their surroundings. Their drawings are, for this reason, detailed and expressive (even if they are badly executed). An adult, on the other hand, rarely draws, and when he does, the things depicted are stylized and simplified, because the only thing that matters is the communicative value of the object depicted (any picture must be ultimately be showed to someone else to communicate something). For an adult, relating empathically to a carefully drawn subject is judged irrelevant. In one word, the goal of impersonal echoes is to trace a noise back to its intentional source. Consequently, the sound is not a noise any more, but a word that can be introduced by a quotative which has a talking subject. In the same way as two conversation partners must to varying degrees relate empathically to each other in order to guess or interpret each other's intentions, impersonal echoes enable an empathetic connection between a human and an object or animal. Somehow, this connection is useful to "animate" things in narratives and to achieve a better representation of them in the listener's mind. Here is an illustration from a Chad Language named Kera, concerning the brewing of beer:

Siibarj Ss3T) ana kayaqa, asarj sawawa siibarj asai) ba qadti bijigid-bijigidbijigidi.
Kumay osarj bh fe siiba k aqgaf td. Ye harar) asarj ta, hümüg
kaayaw' td, r/dßte kumay wara, kumay ba kammij td, gdßtdrj kumay war'
karap kede td. Ye asa η ba se kumay a, ceererj kumay cer td, ye as a η bh se
kumay α td. Ye sdij kumaykaydi)', karj as ay mdrkdrj war'kap. Hulum paapa
babe cüurü apaya ba, lam-lam-lam addwra karap kede.
'The yeast then, after a while the yeast started brewing BUBBLE-BUBBLE.
The beer started forming foam IDEOPHONE. They sat down
again, took the calebasses, served the beer, the beer of the ancestors,
poured out all of the beer COMPLETELY. They started drinking the
beer; after they had poured out some beer (for the ancestors) on the
ground IDEOPHONE, they drank the beer. As they now drank the
beer, they became quite drunk IDEOPHONE. Nobody kept up his
head, they lay IDEOPHONE COMPLETELY on the ground

Bergman & Dahl 1994, p.414

According the authors,

Narratives appear to be the "canonical" text-type for both ideophones and reduplicated verbs- it appears that the livelier the narrative, the more frequent the ideophones. […] Typically, what happens is that a situation is first described by ordinary linguistic means, and then an ideophone is added to give concretion to the narrative. (p.415)

The narrative is not enough to actively involve listeners in the story. They also must be stimulated by the report of what the brewing beer "said". It is difficult to figure out how the imitation of yeast bubbling has to do with intentions. To explain what I understand by that, I would like to analyze the account of a woman about her reaction to the imminent fall of her building:

The depiction of the sound (*crac crac crac*) the woman heard has only one goal: highlighting the fact that she wonders what it is, that she is engaged in an interpretive process (= a *what-do-you-mean* attitude). "Crac crac crac" means something that should potentially be interpreted. The particularity of impersonal echoes is that they put in the foreground an expressive meaning not for the purpose of meaning negotiation, and consequently point of view alignment, but of recognizing that something *has* an expressive meaning, without that this recognition leads to an active interpretation of what is meant (a falling building, in normal conditions, cannot be a conversation partner). When one is confronted with something that *expresses* something that cannot be *interpreted*, the only way to engage intersubjectively with that thing is to reproduce exactly what it says. In other words, what is expressed is not interpreted, but directly embodied, felt, and reproduced exactly as it is. Exact reproductions are the only way to engage dialogically with something whose expressive meaning is not interpretable. Therefore, in narratives, onomatopoeia are a good way to let the listener feel "in flesh" the story events.

It would be a mistake to believe that the only kind of impersonal echoes in narratives are ideophonic. Symbolic verbs (which are in principle descriptive) can also replace them, such as in the image below:

(from comic series *Lanfeust*)

"Mâche" (Eng. "chew") is not an onomatopoeia nor an ideophone, but an impersonal verb (the potential subject is *ça*, "it"= *ça* mâche, *ça* mache). By a mere convention ("mâche" lies outside a speech bubble), the descriptive "mâche" becomes depictive of the ongoing action. The RE contributes also to the belief that "mâche" is an onomatopoeia, and not an utterance.

These primarily observations dispel the misconception that impersonal echoes form a limited category. Verbs can, without presenting ideophony, be entitled to the same status, and be defined, as any onomatopoeia, as "the naming of a thing or action by a vocal imitation of the sound associated with it (such as *buzz, hiss*)"[17]. Repeating a symbolic word changes it into an onomatopoeia (*Chew Chew*), as well as using conventional signs, e.g. descriptions of smileys are pinched between asterisks to make them depictive of attitudes: *smile*, *sigh*. More generally, somebody saying *and I was like "What?"/ "Huh? / "Oops"/ "OMG!"* produces an onomatopoeia, that we could rename at this point "expressive markers". Onomatopoeia and expressive markers could be different in that that latter is apparently personal, while the former is impersonal (with apparent exceptions such as "Yuk!", which is related to a person in the speech event). Any research on the internet shows nevertheless that expressive markers and onomatopoeia fall into the same category:

(from dreamstime.com)

---

17 https://www.merriam-webster.com/dictionary/onomatopoeia, (accessed on April 2, 2020)

Expressive markers such as "Yuk!", "Yikes!", "Yum!", "Crap!", "Damn it!" engage the speaker in a narrative, whereby his feelings are externalized in order for them to be shared with others. Inner states are understood by taking perspective on oneself. Externalizing emotions is the result of the objectification of one's own inner states. There is therefore no clear difference between expressing the noise an object makes and wording an emotion. Emotions (=affects put into words) belong to "narrative selves" as much as ideophones help to narrate objects and animals in order to bring them into life. They offer an access into the emotional life of somebody or something, and to their identity, i.e. they lead to the recognition of others as "others" with whom anybody can dialogue with. Ideophones invite the listener of the narrative to *interpret* the intentions of a particular (adopting a *what do you mean?* -attitude), and not only to understand what it is passively reacting to. Impersonal echoes do not boil down to ideophones however: RE also offer a productive means to express emotions: "The sky is blue, blue…", "It happened a long, long time ago,…", "They are so, so many things I want to tell you.", etc.

It would be relevant at this stage to list the criteria differentiating personal echoes and impersonal ones:

Personal echoes	Impersonal echoes
In conversations	In narratives
Present the speech of participants	Represent inner states
No sound symbolism	Without or with sound symbolism
obligatory RE	optional RE

Some impersonal echoes such as ideophones cannot be either integrated in the sentence syntactically (they must be introduced by a quotative or occur independently), what makes them at first glance look closer to RS. Their status remains however ambiguous, because they can also be repeated (*yum* or *yum-yum*). Calling them "echoes" in the same way as a truly repeated echo (*The sky is blue, blue*) is therefore partially justified.

Another category of impersonal echoes is typically found in Motherese. It concerns "musical" RD (including rhyming RD, e.g. itsy-bitsy, teenie-weenie; ablaut RD: zig-zag, tick-tack) or hypocoristic ones (name doubling: *Jon-Jon*) conjuring up a sense of familiarity, of play, and endearment.

All these cases of RD are poetic in nature: they "sing" the speech, on the contrary to personal echoes and diagrammatic RD (i.e. repetitions denoting intensity, amplification, specificity and

diversity, vagueness; see next chapter). The sound shift is used for the sake of melody and rhythm. Musicality is crucial for the young child's access to intersubjectivity. Musical RD aims at arousing motricity. A clock makes *tick-tock,* and not *tick-tick,* because the former not only indicates a pulsation, as does *tick-tick,* but also a *rhythm.* A pulsation refers to an underlying isochrony (clapping in the hands during a concert), while a rhythm is a regular succession of accentuated and unaccented notes. In Ablaut RD, the swinging between accentuated and non-accented manifests itself in the alternation of high front vowels [i] and [u] and low vowel [a], or front vowels with middle back vowels [o] which maximizes the pronunciation distance (Miyeon 2005), e.g. Eng. Riff-raff, pincum-pancum, zig-zag, bing-bang, mish-mash, chit-chat, pitter-patter, criss-cross, hippety-hoppety, Fr. prêchi-prêcha, mic-mac, fait ci fait ça, etc.

In rhyming RD, there is a contrast of obstruency, i.e. the onset consonant of the reduplicant is a stop sound, e.g. fender-bender, lovey-dovey, hanky-panky, fuddy-duddy, hoity-toity, hodge-podge, etc. the French equivalent of itsy-bitsy, "fastoche", without being reduplicated, contains both maximal obstruency (*f*as-*t*oche), maximal distance (f*a*s-*to*che), and quasi-rhyming sibilants (fas-to*che*). Maximizing obstruency and distance produces a motor, swinging effect that activates the body in the same way as a rhythm makes us move. It has been proven since a long time that the child access to intersubjectivity through rhythm and musicality (Guerra 2018; Gratier 2009), Neuroimagery has also showed that imagination does not consist uniquely of representations, but also of ideomotor reflexes by which we can engage and interact with the world through mental stimulation (Stépanoff 2019). I have underscored the importance of body images in speech enactments, and it is likely that being subjected to a rhythm is a way for the baby to become intuitively aware of his body, without necessarily objectifying it. Indeed, one can be unconsciously subjected to a rhythm (stamping the ground during a concert is an action that is not necessarily conscious). Rhythm creates intuition that we are moved by something that directly acts on our body. Predicting the rhythm and acting on the source of it is a way to regain control on the way "we are moved". It will not be surprising that nursery rhymes (or other rhythmic plays, such a peekaboo) are a gateway to intersubjectivity, the baby or young child actively interacting to regulate, predict, or adapt the rhythm. If he sees that he can regulate the rhythm which affects him by cooperating with the source of it (the adult), it brings him a positive feeling of body control. Ablaut and rhyming RD are a part of a much more pervading phenomenon which relates to the importance of rhythm for the child in his access to the recognition of others.

Lastly, hypocoristic RE appeal to the "emotional (or expressive) self" rather than to the "communicational self" of the addressee. As other impersonal echoes, they put in the foreground the identity of the target of the vocative, the fact that he is a particular "other". Calling somebody

"Lily" or "Lulu", it is acknowledging that he has an identity exclusive to the relationship between him and the speaker, i.e. it is an ontology: Jon is a *brother*, a *boyfriend*, a *son,* etc. Each of these identities is exclusive (Jon cannot be the boyfriend of more than one individual). An (often reduplicated) hypocorism (re)identifies somebody in the same way as an ideophone individuates an object by giving it intentions or mental states. In short: "Jon!" is a call, while "Jon-Jon!" performatively establishes a relationship inside of which everyone is exclusively defined according to each other's particular perspective (brother/sister, boyfriend/girlfriend, father/son,…).

In conclusion, impersonal echoes are not related to collaborating participants, precisely because they do not collaborate, but synchronize around a narrative in which things, animals, humans, express themselves by sounds that are understood as the wording of inner emotions and experiences, and with which one can empathize. In short, the role of impersonal echoes is to recognize each other's communicative intentions.

3.5. Conclusion: Echo as positioning on other's assessments

In the introduction, I have said that all forms of RP are markers of speech enactment, viz. linguistic resources enabling negotiation of expressive meaning. In the case of RS and partially, of RE, the negotiation is unilateral: one replicates in order to steer the inference. But when it comes to echoes, everything becomes more intricate. By uttering echoes, two partners engage much more collaboratively in the negotiation of expressive meaning. Let's take the four of them one by one:

• Simple echoes: The listener insert intra-turn units aiming at steering the conversation in a coordinated way (the Japanese term for back-channeling, *aizuchi*, "mutual hammering", renders very well this idea of how a dialogue is forged interpersonally through short responses of the listener[18]). Simple echoes are listener's "interferences" directing how the actual speaker has to (re)interpret what he says; it is a feedback that modulates the way the speaker understands himself. Simple echoes distribute the interpretive process between two agents in a joint act of participation. It should be noted that simple echoes must not be conflated with non-verbal back-channels, the latter having only the role of validating or invalidating the content of the speech, without interpenetrating with it.

• Reflexive echoes: This time, the speaker regulates himself the space of negotiation that he has opened. This may be due to the lack of responsiveness of the listener, or simply the fear to

---

18 Drawn from the wiktionary entry:
   https://en.wiktionary.org/wiki/%E3%81%82%E3%81%84%E3%81%A5%E3%81%A1#Japanese (accessed on May 6, 2020).

engage with the other. The speaker splits up and becomes successively the producer and the interpreter of his own speech. This process of auto-negotiation may or may be not validated by the listener, who has in consequence a reduced role. It does not prevent that reflexive echoes are uttered by two persons at the same time (adjacent reflexive echo), what in turn does not prevent that the adjacent reflexive echo can be realized by one person (mono-adjacent reflexive echo).

- Recursive echoes: recursive echoes derive from a mirroring effect, in which somebody copies words from somebody else, before sending them back on the original producer. The producer of a recursive echo wants to show that he does not want to participate in the negotiation of meaning, because he disagrees or is indifferent to the partner, his attitude, his intention, his emotions, etc. The expressive meaning that must be negotiated is rejected.

- Impersonal echoes: finally, two partners may just decide to synchronize around a narrative, a common imagery, in which emotions, identities, desire to communicate, etc. are brought to the surface. Impersonal echoes are a way to recognize the *subjectivity* specific to something or somebody. This is a purely a relational and empathetic process. Ideophones, hypocoristic RE, ablaut and rhyming RD all contribute to interpersonal recognition, recognition of feelings and sensibility, of individuality, humanness, etc. Narratives are a preliminary stage to dialogue, as partners tend less to communicate than relate to each other and to co-define themselves. Impersonal echoes are a safe way the share a common perspective on things, to understand them under the same point of view. Because of this, it cannot be spoken of "negotiation" of expressive meaning. Expressive meaning is not negotiated, only shared at an embodied level.

In the next chapter, I will try to know if this categorization can apply to RD, i.e. entrenched form of RE. If this not the case, a new mode of regulation of expressive meaning in RP could be found that relates (almost?) exclusively to RD.

4. Reduplication

Montaut addressed in her article "Reduplication and echo words in Hindi/Urdu" (2009) three kinds of RD based on Culioli's enunciative theory, which I propose to match with my own typology of echoes, with the inclusion of standard RE which constrain inferences: Total RD, Partial RD, and Contrastive Focus RD. After summarizing the content of the article, I will see if eventual correspondences can be made. To better account for non-fitting cases, I will seize Montaut's concept of *mode of presence* and elaborate on it. The tension modulation provoked by a RD (e.g.

intensity, attenuation, extension in time, etc.) will be assumed to stem from the act of sharing presence (i.e. sharing a perceptual viewpoint on something). Lastly, tension modulation will be compared with aspectuality. This last point will only aim at sketching avenues for future thinking about a possible link between grammar and speech enactments.

4.1. Classification

According to Montaut, total RD is set up by a modification of "the scheme of individuation of the notion" (p.22). A notion corresponds to a qualitative representation of a thing, a category (e.g. a dog) and a scheme of individuation to a set of extensions of this notion (dogs, the dog, a dog,…). Non-centering of a notion's occurrence through operations as scanning is the main characteristic of total RD. Scanning "means that you have to scan the whole domain without finding a possible stable location ("any dog", "which dog") (p.22). Intensiveness, distributive meaning, iteration of action are nuances added by RD due to the impossibility to locate a precise occurrence.

Contrastive focus RD, in the case of plural nouns, also involves the scheme of individuation of a notion:

a *yahâN   mahilâeN-mahilâeN   baiTheNgî*
   here   women-women   will-seat
   'here only women / women and only women will seat'
   (context: there are too rooms, one for men, one for ladies)

b *bookmarkoN- bukmârkoN meN hî   bât   hotî calî gaî*
   bookmarks-bookmarks in just   speech be went
   'the conversation went on exclusively by means of bookmarks"
   (two lovers strictly looked after by the girl's family: M. Joshi, K)

   p.26

"Women-women" and "bookmarks-bookmarks" are only meaningful in a context where they are opposed to, respectively, "men" and "other communicative devices". RD tends to homogenize a concept by disqualifying a second option (thus not women-men, but women-women.). Contrastive focus RD can be found in adjectives too, e.g.

a *yah rahâ tumhârâ kok, pî lo,   ThaNDâ-ThaNDâ hai*
   here is   your coke,   drink take, cold-cold is
   'here is your coke, drink it, it is nicely cold' (not "very cold")

b *yah lo   tumhârî cây. Garam -garam hai,   piyo*
   this take your tea   hot-hot is,   drink
   'take your tea. It is nicely hot, drink it'

The author points out that contrastive reduplicated adjectives never measure a degree (they are not *descriptive*, "very hot", "very cold"…) but a subjective appreciation (… but *qualitative*, "nicely hot", "nicely cold") that "neutralizes the feature 'differential' in the property in order to emphasize its manifestation: conformity with the typical idea of a good tea (nicely hot, but precisely not too hot) sets the value, shared par the subject S and hearer as imagined by S" (p.33). In contrast to contrastive reduplicated adjectives, which convey an idea of *saturation* (i.e. an optimal degree), other reduplicated adjectives can be associated to direct perception and refer to a modality ("a particular stand of the speaker", p.35) and denote *approximation*:

a   *nice   nîlâ-nîlâ pânî*   *bahâ rahâ thâ*
    below blue-blue water   flow PROG IMPFT
    'bluish water was running below'

b   *sâmne nîle-nîle pahâr kî rekhâ*   *dikhâî de rahî thî*
    in-front blue-blue montains of line   be-seen PROG IMPFT
    'the bluish line of the mountains could be seen in front'

"blue-blue mountains" cannot be paraphrased "the truly blue mountains" or "the nicely blue mountains" as in the examples above, but "bluish". The attenuative meaning is related to a direct perceptual experience, and therefore often co-occurs with verbs of perception ("seen" in *b*). In either case, approximation or saturation, both meanings result from a "neutralization of the differential property in the adjective" (p.36).

The article deals next with *Partial reduplication*. Partial RD extends the notional domain, or de-centres it (p.38). In Hindu, it is mostly done by the *v-* alteration (structurally similar to the *shm-* in English, see next section), as in *pen-ven* ("pen and the like"), *shâdi-vâdi* ("marriage and son on") (p.38). Partial RD includes the neighboring zones of the notion (i.e. what does not belong to the notion, but is part of its semantic field, for instance "pen and the like" = "pen" and everything used to write.) V-RD is often combined with parodic, polemical or derogatory aims (p.42), as in:

a   I understand you very well, I too have been young and communist
    lekin kuch din <u>krânti-vrânti</u> kî bât mat kîjie,
    but some days revolution-echo of speech NEG do
    krânti (*vrânti) kî hî khâtir ke lie
    revolution (*echo) of only interest for
    'but forget for sometimes your <u>revolutionary big talk</u> (do not speak of <u>revolution-echo</u>), in the interest of revolution itself (echo is impossible)"

The reduplicant "re-qualifies" the simplex, which is systematically taken from a previous turn. The positive viewpoint conveyed by the first speaker about revolution is negatively re-qualified by the second speaker (*krânti-vrânti*). Two conflicting viewpoints are present in partial RD: the positive one, which is initially made explicit, and the negative one, which rejects it (p.46). This confrontation is not necessarily polemical; it can also mean that each one has divergent opinions of equal value on the matter that need to be respected. Non-centering of the notional domain allows a diversification (or a blurring) of the notion that can be sometimes interpreted as derogative, sometimes as accommodating.

The article that I have summarized has two advantages:

•	Montaut stands up against the general belief that RD is an icon of excess, following the principle "Two is more". On the contrary, RD follows systematic rules (modification of the scheme of individuation of a notion; de-centering of a notion). "Far from being the icon of excess […] [reduplication] operates in a systematic way." (p.55)

•	Repetition operates on an *intersubjective* level. It allows adjustments between speakers: "Modifying the scheme of individuation may involve the subjective interaction of the speaker and hearer's viewpoints. Partial or alliterating reduplication […] involves almost always the confrontation of two distinct viewpoints. (p.55).

It still remains to find to what extent the classification proposed by Montaut matches with my own, and whether they are discrepancies between them or not.

4.2. Echoes in RD?

The second category Montaut addressed in her article, *Partial RD*, is very much in line with what I have called *recursive echoes,* the main difference being that echoes are not subjected to productive rules such as the v-alteration, which de-centers a notion to include in it similar things ("and all that stuff"), nor to rhyming RD found in lexicalized partial RD such as Finnish "hölynpöly" (nonsense), Eng. hodgepodge, claptrap, etc., which most of the time refer to the speech itself, and not to a word in particular. Lexicalized partial RD are besides not always rhymed, e.g. Fr. Bla-bla. It would be therefore dangerous to systematize these general tendencies. Nonetheless, the negative connotation associated with Partial RD is very salient and has to be taken into account.

I propose to compare reflexive echoes with a more familiar type of partial RD than v-alteration, namely the *Shm-RD*. Here are two examples:

A: He's just a baby!
B:"Baby-shmaby". He's already 5 years old!

A: What a sale!
B:"Sale, schmale". I'm waiting for a larger discount.[19]

First, let us start by calling the attention on the telling French translation of *shm-*, "tu parles": "Ce n'est qu'un enfant ! - *Tu parles*, il a déjà cinq ans"/ "Quel offre! - *Tu parles*, je vais attendre une réduction plus conséquente". The speaker repeats what the other said, but at the same time rejects it by reflecting it to the original author instead of integrating it into his own speech. The onset consonant is a way for B to explicitly let know that he animates the word repeated without embracing it. The reduplicant includes the necessary change that will show that the echoing person does not endorse the content of the word, but reflects it to the original author ("Hear yourself talking!"). In the case of RD, the difference between the voices is grammaticalized: the simplex is the original, the altered reduplicant the dismissive echo. So, *shm-* RD always reacts to an utterance in order to (gently or not) discredit the offender by taking overtones:

1. "The doctor says he has a serious virus? Virus, shmirus, as long as he's healthy."

2. "Who said that? Fred? Fred, Shmed, what does he know?"

3. "The psychiatrist says he has an Oedipus complex. Oedipus, Shmoedipus, so long as he loves his mother."[20]

In 1. and 3., the speaker disdainfully dismisses the reported assessment of the doctor and the psychiatrist respectively. If for instance "virus" had been repeated without modification of the initial consonant, he would have shown that he potentially believed what the specialist said: "Virus, [he meant:] virus? but that's terrible!". A partial RD indicates however that the speaker rejects the content of the echoed simplex. This can be often offensive for the interlocutor himself: in (2), this is not Fred who is laughed at, but the one who answers to the speaker's question *Who said that?*: "Who said that? - *Answer-* Fred? – Fred, schmed […]!". Fred is of course despised, but indirectly, the interlocutor is also held in contempt, maybe for giving so much credit to Fred's silly words, which he reports uncritically to others. His naivety is thus also criticized by his interlocutor.

Montaut addressed a second category of RD, namely *total RD*, which corresponds roughly to what I have called "inference constraint". The neutralization of the differential feature homogenizes a concept that could be ambiguous in a context in the same way as a RE orients what should be

---

19 Drawn from the wikipedia entry: https://en.wikipedia.org/wiki/Shm-reduplication (accessed on May 2, 2020)
20 Source: "Yiddish, shmiddish–why do we repeat a word but start it with "shm-"?"
https://www.straightdope.com/columns/read/1997/yiddish-shmiddish-why-do-we-repeat-a-word-but-start-it-with-shm/
(Article published on February 12, 2002. Retrieved April 18, 2020).

understood, as Montaut puts it: "the speaker does not *a priori* allows the possibility for the hearer to conceive the entity with another property" (p.33). However, modal reduplicated adjectives (i.e. RD denoting approximation or intensification) fit less well: the degree of saturation and approximation involves more than a guided inference, but the adoption of a "mode of presence", which "relates to a particular stand of the speaker: tell the world (a given entity of the world) such as he is confronted to it" (p.35). For instance, if someone rushes out of a kitchen with a bowl of soup in the hands, saying "hot hot!", he communicates a particular mode of presence. The hotness of the bowl is not just a quality, but the perception that somebody has of this hotness. To understand "hot hot!", the listener has to empathize with the carrier of the bowl in order to *feel* the sensation. RD relating to a mode of presence are numerous, yet do not find their place in one of the echo-categories. Obviously, they are connected to feelings and empathy, what brings them close to impersonal echoes. The problem is that the latter only entails the recognition of feelings, not the sharing of them. A third category must therefore be created, that does not reduce to general RE or echoes, but corresponds to a "mode of presence", which I will call "diagrammatic RD/RE".

4.3. Diagrammatic RD

The adjective "diagrammatic" refers to a kind of RD (as well as RE, such as "hot hot") that can be defined through a graphic representation of the tension level of a particular concept: a RD's meaning can be modulated along two axis: intensity (Y) and extension (X). Extension refers to diffusion, intensity to concentration. Most RD play on an increase and decrease of intensity and extension, for instance Afr. Drie-drie storm deur die hek (lit. three-three storm through the gate, Eng. "Groups of three storm through the gate", Van huyssteen 2001, p.285). The red star represents the simplex used alone, "three", which has an intensity of "three" and no extension. When reduplicated, "three" has an extension in time (three by three):
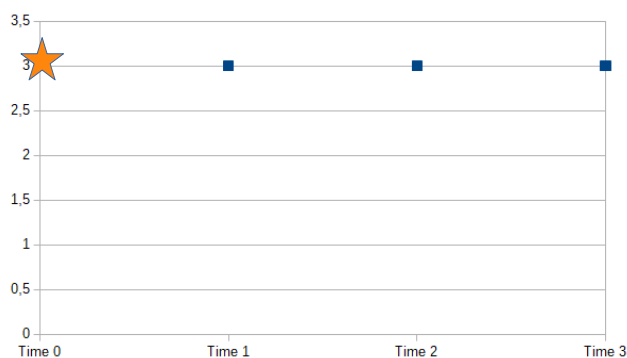
Other examples of the same kind would be Jamaican Creole *luk-luk* ("to look now and again"), Afrikaans *Frans vorder rus-rus* ("Frans progresses while resting from time to time"). Each verb has an intensity (e.g. looking is stronger than *having a glance*, and weaker than *observing*), and through the RD, develops an extension in time or space. But how could a mere RD precisely specifies the degree of saturation and extension? It is indeed true that RD can be polysemous, such as in Jamaican Creole *redi-redi* (red-red; Fischer 2011, p.60), which can mean reddish, very red, or red-spotted. Both intensity and extension are modulated. The star represents the centered notion (the simplex):
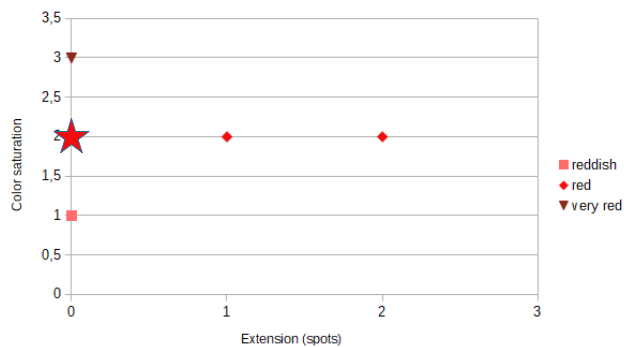
"Very red" increases the intensity, "reddish" decreases it, and "red-spotted" extends it. Tension modulation in RD is found throughout the world: here is a similar case in Lushootseed, a Salish language once spoken in the modern-day Washington State:

17. ʔujúbʔalikʷčəɬ . . . . 'We danced.'
18. ʔujúbjubʔalikʷčəɬ . . . . 'We danced a lot.'
19. ʔujújubʔalikʷčəɬ . . . . 'We danced a little.'
20. ʔujubúbʔalikʷčəɬ . . . . 'We danced all over the place.'

Urbanczyk 2001, p. 125

17. is the simplex, 18. and 19. are reduplicated forms modifying the intensity (increase or decrease), and 20. the extension. A diagrammatic RD potentiates all these "x-tensive" meanings, and must be actualized in context to be interpreted. I hypothesize that a reduplicated form such as this one is not semantically polysemous, but pragmatically ambiguous, where "ambiguity" must be understood as functional: namely, the main function of RD is to downgrade a conventionalized simplex belonging to the second-order languaging to the domain of first-order languaging, i.e. the place where discourse must be collaboratively negotiated. Tension modulation is a way to trigger a process of "finding meaning together". RD makes a word ambiguous in order to activate a disambiguation process achieved mutually on the basis of contextual inferences. "Ambiguity" must

not be understood in terms of Relevance Theory, that is as a trade-off between production and processing effort, or clarity and ease. As Piantadosi & al. (2011) put it: "where context is informative about meaning, unambiguous language is partly redundant with the context and therefore inefficient; and second, ambiguity allows the re-use of words and sounds which are more easily produced or understood" (p. 281). The first hypothesis of Piantadosi makes sense, provided that it does not rule out the possibility that the role of ambiguity is not so much to avoid redundancy with contextually available information, as to create purposely a displaced context (a space of meaning negotiation). The second hypothesis is more difficult to apply, though: saying "reddish/very red/red-spotted" is cognitively as much affordable as reduplicating; logically, no more processing effort is demanded to compose a *reduplicated stem* as to compose a *stem + derivational morpheme (-ish)* or *+ gradient adverb*. Therefore, it can be assumed that the ambiguity of RD is not to strike a balance between clarity and ease, but to collaboratively "stage" the meaning rather than merely communicate it. "Ujubjubalik" makes ambiguous "ujubalik", whose meaning is univocal. The disambiguation process initiated by the speech enactment (the RD) requires that both partners collaboratively conform to one specific meaning.

Of course, one condition for that is that, let's say, the apple that is described (I'm referring here to the first example, *red-red*) should be absent. If I reduplicate "red" with an apple in my hand, this is to draw the attention on it (an exclamative): "*See* how red it is!". I intend thereby to share a specific attitude towards the apple. Clearly, RD has an emotional component which does not boil down to informativeness. Tension modulation cannot be restricted to mere information sharing. Dispositions, attitudes or beliefs about something must be shared to hold value, and consequently include in their linguistic expression involving strategies such as RD. Someone uttering "See how the sky blue is!", or preferably if the sky is absent, "The sky was blue blue! [try to imagine!]" prompts his interlocutor to guess what is the source of his enthusiasm. To do that, the interlocutor has only one choice: raising the head towards the sky (or imagining the sky), and evaluating it qualitatively: is the sky very blue or bluish (intensity), or are there stretches of blue sky regularly to be seen through portions of thick clouds (extension)? Ultimately, the evaluation process aims at interpreting the exclamation or the RD of the partner, not to scientifically determine the blueness of the sky. By attempting to share the perspective of the other and to empathically conform with his feelings, a specific mode of presence is created. Determining the tension of a quality happens by interpreting the source of the exclamation in order to find out the expressive meaning of the utterance. In the present case, the sky is the support of the interpretation. In comparison, In Afr. "Drie-drie storm deur die hek", it would be the gesture the producer makes while speaking, eventual

69

background knowledge, as well as logic (it would be difficult to evaluate "drie-drie" on the axis of intensity) which would orient the interpretation.

Impersonal echoes and diagrammatic RD are sometimes difficult to distinguish. Actually, *the sky is blue blue* can be understood as either one or the other: *the sky is blue blue…* said in a melancholic tone is a means for the speaker to express his emotions. There is no modulation in the color tension. Maybe the blueness of the sky recalls the speaker buried memories or moves him in a certain way. The suspension points further demonstrate that the utterance is an emotional "outpouring" which exclusively relates to an inner state, i.e. nothing is said about the sky itself, because the topic is actually the inner state of the experiencer. In Japanese pragmatic studies, this effect is called *futaku*, lit. "commiting, referring to" (Maynard 2002), that is "a method for expressing one's feelings by borrowing something concrete" (p.107). The role of *futaku* is to arouse empathy, what direct description of feelings ("I am melancholic") does less well. This is maybe due to the fact that "I am melancholic" is a speech act: the utterance has consequently an illocutionary effect (e.g. a request), as well as a potential perlocutory effect (the speaker expects that this implicit request will be answered), for instance: "I am melancholic [so please give me antidepressants]". Such speech act does not build empathy. However, saying to someone else: "The sky is blue blue…" thwarts the illocutionary effect. As the utterance cannot be interpreted anymore (i.e. there is no implicature to infer), it is understood as the projection of a pure desire to communicate ("pure" means that there is in fact no propositional content): the foregrounded meaning is "I express myself". The potential addressee understands that the only way to respond to this utterance is to empathically reproduce the whole behavior of the sender (not only the text, but also the way the words have been produced, as in an onomatopoeia), in order to "align" with the partner, who did not let him identify what he wanted to communicate (= the illocution).

On the contrary, an exclamative sentence like *the sky was blue blue!* aims at sharing presence by a modulation of the notion *blue*. The speaker tries to make the interlocutor visualize the sky as he himself experienced it. In this case, one could assume that the blueness is intensified rather than attenuated.

Finally, let us note that RE (often embedded in narratives) can also be diagrammatic (however, it appears to bee less systematic than in RD). Here is a fragment of a geography course:

(Dynamic earth lecture)
now after the shoreline has moved inland, that particular point is under deeper water, than previously, so we are depositing, sand... [well, imagine, the shoreline being able to move even further inland, in other words that basin is either sinking, or sea level is rising very very rapidly, and therefore this particular unit, being able to form a very very large area, that keeps going on and on and on, for several, hundreds of kilometers, uh inland.] so typically therefore, (…)

Two locutions (in red) frame the narrative. Three repetitions (in green) occur inside it. At the end, the "typically" signals that the lecturer makes the switch between a *token* (the narrative) and a *type*. *On and one and on* is typically found in such narrative with illustrative purposes: the professor makes a little digression to illustrate his point, only to let his students generalize the example. *On and on and on* is a hinge that links the particular to the typical (*typically*). The teacher wants his students to imagine every concrete situation on basis of his own specific example, so that they can reach the abstract level. While the two *very* are intensified, *going on* is extended. Tension modulation is thus not exclusive to RD.

4.4. RD and Aspectuality

There are still cases the diagram does not account for. In spoken French, it may happen in familiar register to answer to a question such as "Where are the biscuits I have put on the table" with "ils sont mangés mangés" (lit. They are "eaten eaten" = "completely eaten"). The RD denotes completeness and can be linked in this regard to the perfective aspect. Olga Fischer (2011) has observed for that matter that the semantic extent of RD matches exactly the one of the Dutch/German perfective prefix *ge-*. In order to show the connection between the two, she lists all the functions endorsed by this prefix in modern Dutch/German. I reproduce hereunder her enumeration:

- generally in past participles (result, perfective)
- in count nouns expressing result (often derived from past participles): *gezegde/Gesagte* 'what has been said, a saying', *gewoonte/Gebrauch* 'what has been done before, a custom', *gestalte/Gestalt* 'appearance', *gesprek/Gespräch* 'a talk'
- in collectiva formed from nouns: *gebergte/Gebirge* 'mountains', *Gewolk* 'clouds', *gezusters/Geschwister* 'sisters'/'brother(s) and sister(s)', *getij/Gezeit* 'tides',
- in abstract nouns derived from verbs: *gezang/Gesang* 'singing', *gezicht* 'sight', *gehoor/Gehör* 'hearing' (in Dutch also 'audience'), *gezag/Gewalt* 'authority' (from Dutch *zeggen* 'to say'), *gewicht* 'weight, importance', *geweld/Gewalt* 'force'
- in instrumental nouns: *gezant/Gesandte(r)* 'messenger' (from *zenden/senden* 'to send'), *gewicht/Gewicht* 'a weight' (piece of metal used in weighing), *getouw* 'loom' (from *touwen* 'to work on'), *Gebläse* 'blow-instrument, compressor', *Gefährte* 'companion'

- in adjectives expressing a result: *genegen/geneigt* 'inclined'(from *nijgen/neigen*), *geschikt* 'apt', *geletterd* 'lettered', 'educated', *geleerd/gelehrt* 'learned', *geruit/gewürfelt* 'checkered', *gelakt* 'varnished'

- with a pejorative sense: *gepeupel* 'mob', *gespuis/Gesindel* 'rabble, scum', *gedoe/Getue* 'a to-do', *geklets/Geschwätz* 'twaddle', *gelazer* 'bother', *gezeur/Getratsch* 'drivel, bothersome behaviour'

(Fischer 2011, p. 67)

Fischer only examines the modern functions of *ge-*, but this prefix had other functions in Middle High German (possibly in Diets). In this language, present forms augmented with an intensive *ge-* could have a future meaning:

mhd.: *ich weiz vil wol, waz Kriemhild mit diesem schatze* **getuot**
nhd.: Ich weiß sehr gut, was Kriemhild mit diesem Schatz tun **wird**.   Nibelungenlied v. 1272
Eng: I know very well to what end Kriemhild **will** use this treasure (my own transl.)

It also expressed plus-quam-perfect, e.g. *gelebete* in MHG. could mean "gelebt hatten" (Eng. "had lived"). From these two older functions, it can be observed that *ge-* and RD have much more in common than perfectivity. Future and Plus-quam-perfect are an "increase" in something with respectively present and perfect as starting point on the axis of intensity. Aspectuality and tension should therefore not be considered individually. On the basis of Fischer's enumeration, I provide a list of functions in RD showing similarities with *ge-*:

- RD indicates perfectivity: Jamaican *kata*, "to scatter" > *kata-kata* "scattered". Ndjuka *dee*, "to (be) dry, *dee-dee*, to be (in a) dry (state) (Fischer 2011, p.7)

- (partial) RD indicates future: tagalog *tawag*, "call"> *tatawag*, "will call", *sulat*, "write"> *susulat*, "will walk" (Kauffmann 2015)

- (Internal) RD indicates plural, just as *ge-* collectiva: Indonesian *kapal*, "ship"> *kapalkapal*, "ships", Samoan *le tamaloa*, "the man"> *tamaloloa*, "the men" (Kauffman 2015)

- RD can be pejorative just as *ge-* prefixed nouns, e.g. v-p alternation in Hindu, Shm- RD, Fr. Blabla ("jibber-jabber"), Finnish höpö-höpö ("nonsense").

- Instrumental nouns: Saramaccan *tai*, "to tie">*ta-tai*, "string", Berbice Dutch, *bain*, "to cover"> *bain-bain*, "lid, cover" (Fischer 2011, p. 6)

- Abstract nouns: Papiamentu *tembla* "to shiver"> *tembla-tembla,* "(the) shivering"

- Result: Sranan *koti* "to cut> *koti-koti* "slice"

Nearly all of them could be related to perfectivity (result, abstract nouns) and intensity (plural, future), The pejorative sense, that we already met in discussing recursive echoes, is also common to *ge-* and RD. Only the metonymic shift to "instrumental" seems at first glance not to fit in either category. However, it leads nowhere to try to connect every point covered by *ge-* and RD in a systematic way. I speculate that RD is above all a way to enact the speech, and to enter a process of meaning negotiation. It is thus not the result, but the process of meaning-making which is important. As I do not have the tools to verify it, I will let this question open. Nonetheless, I would like to illustrate my point with a personal experience: I'm used since a long time to refer to a "perforatrice" (a hole punch) with the word "troutrouteuse" (the word *trou,* "hole", is repeated and completed by the agent noun suffix *-euse,* "-er"), because I've got difficulty in remembering it. The reduplicated action (making holes) describes the instrument (the hole punch). However, I have never used this redesigned version of "perforatrice" without actively searching the affiliation of my interlocutor by showing insistently that I mean something that I cannot express. What is more, the word itself is not enough; it rather goes with the mimic of performing the action. The RD does not "cognitively" relates action and instrument, it is only a superficial way to show that meaning must be negotiated in the context in which it is used, because linguistic resources fall short to describe the notion. If my interlocutor would have caught my meaning, only to confess thereafter that he did not know the word either, then maybe the word "troutrouteuse" would have replaced the more difficult "perforatrice" in each interaction involving the two of us. Put briefly, the RD is just a metasignal that the speech is or must be enacted. The active negotiation takes place on a first-order, multimodal level. "Troutrouteuse" could be a paraphrase of: "The [X] that makes holes". The [X] is an open slot that must be filled in context. This works too for resultatives, e.g. Sranan *koti-koti* (slice) from *koti* (cut) is tantamount to "the [X] that is cut". More generally, perfective RD is the remnant of a meaning that was once situational. Gradually, this device, whose only purpose is to shift from second-order to first-order languaging in cases such as when one forgets a word, becomes lexicalized. It comes to signify lexemes or grammatical notions that are difficult to express because they do not hold (yet or by nature) any referent: states, aspects, instruments, etc.

As showed above, the functions of the prefix *ge-* are not only perfective, but also tensive, just as in RD. Tension modulation and aspectuality are maybe close-related. This remark becomes salient once we look at how aspectuality, tension, and RE are interconnected in sign language. Here is a list, that I draw again from Fischer's article:

73

- slow repetition (this expresses continuity in atelic verbs; or iterative aspect in punctual verbs)

- fast repetition in combination with punctual verbs (this expresses habitual aspect)

- tense hand movement: a long, tense start of the sign, fast progression with abrupt end (this conveys intensive aspect)

- tense hand movement that starts slowly and then speeds up to a long drawn out end (indicating resultative/completive aspect)

- lax hand movement and reduction of the repetition (conveying approximation)

- repetition of a sign for a verb in a contained manner (i.e. the repeated movement is smaller than that of the first movement (this leads to nominalization, e.g. *to sweep > broom, to sit > chair, to compare > comparison, to read> reading*), indicating instruments as well as abstract nouns)

- repetition is sometimes used to express plural by means of a replacement of the movement in space in order to express a scattered plural (e.g. sign for *car > car car* 'cars here and there')

(Fischer 2011, p.66)

Tension of the movements indicate, among others, aspects. In language, RD is an important to modulate tension, because it signals that communication is embodied and negotiated in a multimodal way. RD is not tensive in itself, but is an index pointing to an embodied situation of communication where such tension is made visible, through verbal and non-verbal means. RD is not *iconic* or *tensive* in itself, it only points out that the body comes into play. By contrast, *true* verbal tension modulation can be expressed by submorphemes (Bottineau 2013): specific verbs of actions can contain iconic initial consonant clusters called submorphemes, e.g. *gl-* in verbs denoting light like *gleam, glisten, glow, glare, glitter*, or *kn-* associated with protrusion in *knee, knoll, knob, knead, knock or knell* (Argoud 2010). Submorphemic verbs are, in my opinion, different from verbs of high generality (ex. go, shine, sing), because they rest on tension. It thus means that they indicate a degree on a scale (it shines *more* or *less*, the light is *diffuse* or *concentrated*; *continuous, interrupted, sporadic*, etc.). If submorphemes and RD can be seen as speech enactments, the difference between the two is that the former is an integral part of the embodied enactment, while the latter indicates it.

As a philosophical footnote, tension modulation could be compared to the interpretation of a piece of music. The musical notation refers most of the time to variation in intensity (piano, forte, accent, crescendo, …) or extension (staccato, tenuto, allegreto, fermata,…). The musician animates his performance by modifying the isochrony or varying the intensity, what infuses the piece with emotions:

> Pour qu'un morceau de musique soit expressif et "émouvant", l'interprète du morceau doit s'écarter, par moments, du rythme établi, du temps marqué par le métronome […] le rythme expressif permet aux musiciens d'introduire leurs intentions ou leurs idées dans le flux musical (Gratier 2009, p. 36)

Furthermore, the philosopher Susanne K. Langer also observed that

> The tonal structure we call "music" bears a close logical similarity to the forms of human feeling – forms of growth and attenuation, flowing and stowing, conflict and resolution, speed, arrest, terrific excitement, calm or subtle activation and dreamy lapses" (Langer 1953, p.27)

The musician constantly shares his presence by means of tension modulation. In other words, he relates his pre-reflective experience of the world (the way the environment naturally affects his body) to the others. Hence, the affect can be potentially embodied (co-felt) by an audience, or reflexively by the musician himself. In the process, affects are transformed into emotions. I'm prone to think that RD and RE resting on the principle of *presence sharing* have an emotional component that should not be disregarded.

5. General Conclusion

In this work, the main focus came to lie on dialogue. It would be oversimplifying to say that a dialogue is a transfer of successive signals encapsulated in "turns". Signal and dialogue are two types of communication that do not converge. Every utterance is dialogic, because the partner (virtual or not) has an effective role in the way each utterance unfolds. At all times, a common interpretive stance must be maintained in order for the communication to succeed. The syntagmatic axis is for this purpose often disrupted by instances of replication. Forms of replication do not operate on a grammatical level, but on a dialogic one: it actively regulates how interlocutors interpret each other, or ensures that interpretation is still possible. When the goal of the speech is not to transfer information, but to negotiate expressive meaning in order to share a perspective, the speech is enacted. All kinds of replication, to varying degrees, signal that the speech is enacted, or that the speaker who replicates looks for the speech to be enacted. The negotiation can be carried out along three procedures: (i) constraining inference, (ii) echoing, (iii) sharing presence.

Some kinds of reported speech and repetition constrain the inference by ruling out interfering interpretations. This unilateral regulation of expressive meaning is limited, though. Often, the

partner is put to contribution in the process through the use of special category of repetition (including some rarest cases of reported speech) called *echoes*, which break up into four branches: *simple echoes* are a sign that the listener effectively contributes to the way the speaker interprets is own speech; *reflexive echoes* show that the speaker regulates the expressive meaning alone by splitting up into producer and interpreter; *recursive echoes* indicate that one of the participants do not want to affiliate with the communication intentions of his interlocutor; *impersonal echoes* form a particular category in that it draws on the recognition that somebody or something expresses meaning, viz. that he has an identity (he is an other) and can engage in communication. Due to the emotional potential of the latter, it is not surprising that impersonal echoes are found in Motherese, narratives, or intimate speech. *Recognition* replaces here *negotiation*.

If impersonal echoes comprise some cases of reduplication (ablaut, rhyming), most reduplicated words need another paradigm than *echoes* to be comprehensible. I have thus introduced the concept of *mode of presence*, which Montaut (2009) applied to reduplication. After summarizing her article, I borrowed her categorization of reduplicated words (total, partial, and contrastive) and compared it with echoes and repetition. Partial reduplication was similar to recursive echoes, and some cases of contrastive focus reduplication to inference constraint, but it was also obvious that some reduplicated adjectives, which relate to a particular mode of presence, did not fit in any of the concepts introduced so far. To compensate for that shortcoming, chapter 4. addressed what I have called "diagrammatic reduplication", i.e. instances of reduplication that can be projected on to a coordinate system according to their level of intensity and extension. Modulation on these two axis results in the sharing of a mode of presence, i.e. a qualitative perception of something. Aspectual reduplication (perfectivity in particular) is not as disconnected from diagrammatic reduplication as it seems at first glance. In sign language, a specific aspect is conveyed through tension variation. It is thus possible to understand aspectuality through the lens of tension modulation, and hence, through diagrammatic reduplication. In this regard, the perfective germanic *ge-* displays a set of meanings that have a lot to do with intensity (collectiva, future), what points to the fact that the border that separate aspect and tension modulation is thin. Reduplication, however, does not directly affect tension; it is only a signal that the speech is enacted (downgraded to first-order languaging), and that meaning must be negotiated on the very place of the interaction. Once the dialogue becomes fully embodied (multimodal), other means are (supposedly) available for modulation to happen. By contrast, submorphemes determine directly the projection on the diagram.

In this thesis, I have wished to highlight the permanent negotiation of expressive meaning between participants in specific contexts. Meaning is not only understood, it is also interpreted, and this requires the involvement of all interlocutors. There is definitely an emotional component that has to be taken into account: basically, speech participants have to reciprocally conform to each other's attitude, to build a situational, co-determining relationship on the very moment of the interaction where each one can recognize in the other his specific attitude towards the topic, the way he is confronted to it. This process is not necessarily collaborative; one may also impose to the other his way of seeing, or, on the contrary, accommodates to the way of seeing of the partner. Replication, as marker of speech enactment (contextual meaning negotiation), offers an access to the study of dialogue dynamics, and hence to our life experience overall, as Fuchs (2017) puts it:

> Our experience of the world is not a solitary achievement, but is based on a continuous intersubjective co-creation of meaning, or *we-intentionality*. We live in a shared lifeworld because we continuously create or "enact" it through our coordinated activities and "participatory sense-making" […]. This includes circular processes of mutual understanding, negotiation of intentions, alignment of perspectives, and reciprocal correction of perceptions - processes that take place in every interaction and communication with others. Thus, intersubjectivity implies a co-construction of meaning through mutual interpretation and perspective-taking. (p.93)

I hoped that I have been able to decipher some of the linguistic evidence that demonstrates this permanent "co-creation of meaning" in our everyday interaction.

Transcription conventions

For the extracts drawn from Holt and Clift (eds.) 2007.

Acknowledgements

Transcription conventions

The method of transcription used within this volume was developed by Gail Jefferson (but see the Appendix to Chapter 6 for some additional notations used within that chapter.) The system attempts to capture some of the features of the interaction relevant to its organisation, including turns, overlapping talk, pauses, and intonational features such as emphasis and marked rises and falls in intonation.

This summary of transcription notations relies heavily on Atkinson and Heritage (1984) and ten Have and Psathas (1995).

Simultaneous turns

Where turns are begun simultaneously they are marked by a single left-hand bracket at the start of the turns:

    Tom: [I used to smoke a lot when I was young
    Bob: [I used to smoke Camels

Overlapping turns

When turns do not start simultaneously, the point where the new turn begins is marked by a left-hand bracket within the first turn and at the start of the new turn:

    Tom: I used to smoke [a lot
    Bob: [He thinks he's real tough

Where overlapping talk stops, it is marked by right-hand brackets:

] Tom: I used to smoke [a lot] more than this
] Bob: [I see]

Contiguous turns

When turns are latched (i.e. there is no interval between the end of one turn and the beginning of a next) equals signs are used at the end of the first turn and the beginning of the second:

= Tom: I used to smoke a lot=
= Bob: =He thinks he's real tough

Equals signs are also used where a transcriber has put on to a new line elements of a turn that form part of the continuous flow of speech often due to intervening overlapping talk:

= Tom: I used to smoke [a lot more than this=
  Bob: [You used to smoke
= Tom: =but I never inhaled the smoke

When overlapping turns end simultaneously and a subsequent turn is begun without an interval, the overlapping turns are followed by right-hand brackets and equal signs. The new turn is marked by a left-hand bracket at the beginning:

]= Tom: I used to smoke [a lot]=
]= Bob: [I see]=
= Ann: =So did I

Intervals within and between turns

Intervals in talk are timed to the tenth of a second and marked by numbers in parentheses either within a turn at talk:

(0.0) Lil: When I was (0.6) oh nine or ten

or between turns:

       Hal: Step right up
(0.0)   (1.3)
       Hal: I said step right up

An interval of less than, or around, one tenth of a second is marked by a period within parentheses:

(.) Dee: Umm (.) my mother will be right in

Characteristics of speech delivery

Punctuation is used to convey characteristics of speech delivery rather than grammatical units.

Sound stretch

A colon indicates a stretched sound:

: Ron: What ha:ppened to you

Longer stretches of sound are indicated by multiple colons:

::   Mae: I ju::ss can't come
::: Tim: I'm so::: sorry re:::ally I am

Cut-off

A single hyphen indicates that the prior word or sound is cut off (i.e. abruptly terminated):

- C: Th' U:sac- uh:, sprint car dr- dirt track...

Intonation

A full stop indicates a fall in tone usually at the end of a unit:

. Jenny: They're a lovely family now aren't they.

A comma indicates a continuing intonation:

, Tony: That really makes me ma:d,

A question mark indicates a rise in intonation:

? V: A do:g? enna cat is different.
   P: Yih ever take'er out again?

80

Transcription conventions

↑↓ L: <u>AND</u> uh ↑<u>we were</u> looking rou-nd the ↓sta:lls 'n poking about...

Marked rises or falls in intonation are marked by upward- or downward-pointing arrows immediately before the shift:

Less marked rises or falls in intonation along with some stretching can be marked by underlining immediately preceding a colon for a fall:

a: L: we (.) really didn't have a lot'v cha:nge...

and underlining of a colon for a rise:

a: P: I'm (h) eyeing your cho:p up mi:nd.

An exclamation mark indicates an animated tone:

! C: An that! so what he sez.

Emphasis

Underscoring indicates emphasis:

<u>Still</u> L: ...he said oh: hhello Lesley (.) ↑<u>still</u> trying to buy Something ↑° nothing,

Croaky voice

An asterisk indicates croaky voice in the immediately following talk:

* D: Wul mayb- maybe *uh- uh good thing...

Volume

Increased volume is marked by the use of upper case letters:

PUT D: ↑PUT THE FI:RE ON

Decreases in volume are marked with a degree symbol at the start and end of the quiet talk:

° ° J: ...I'll go ahea:d, and, .hh pay for it when it comes and °he'll never kno:w°,

Aspiration

A series of 'h's preceded by a dot indicates an in-breath:

.hhh L: .hhhhhh Santa Claus.

Out-breath is indicated by a series of 'h's without a preceding dot.

hh J: I'm not sure hh- who it belongs to

Laughter and smile voice

Laughter is indicated by an attempt to convey the sound using 'h's and vowels, while laughter particles within words are indicated by an 'h' or 'h's in parenthesis:

eh D: ↑UH you fi:nished with that ch↑o::p p↓et eehh he he °he [he he he°

(h) P: [I'm (h) eyeing your cho:p up [mi:nd.
     D: [hah hah hah .hhh
     D: it's cool...

Smile voice is indicated by pound or dollar signs immediately before and at the end of the affected talk:

$/£ P: ...und you might $knock a few blocks out of position$

Speed of delivery

Talk which is noticeably quicker than surrounding talk is marked by 'more than' and 'less than' symbols either side of the fast talk:

> < P: >In fact d'yuh think they will< enjoy co:mpany.

Indecipherable sounds

Where a transcriber is unable to make out a sound or a series of sounds, spaces within parenthesis are used to convey the approximate extent of the missing talk:

```
           D:    huh hee it'd [bust
( )        P:                 [(
```

Where the transcriber is able to guess at the words or sounds used, these are included within parentheses:

```
(in)       J:    ↑Santa Claus brought it. (in his sle:d).
```

Where the transcriber is unsure of who made a sound, empty parenthesis may be used instead of a name or initial:

```
( )    ( ):    ( )
```

Verbal descriptions

Double parenthesis is used to enclose a description of the talk or some other phenomenon present during the interaction that the transcriber does not want to convey by attempting to represent the sound. For example, in the following extract an 'f' in double parenthesis marks falsetto intonation:

```
(( ))      J:    ((f)) ↑Da:ddy Mommy:
           Bob:  ((sniffle)) He thinks he's tough
```

Presentation conventions

Horizontal arrows to the left of the transcription indicate a turn that the author wants to call the reader's attention to. The significance of the turn will be explained in the text.

```
→          Vera:→   Well I said tuh Jean how about it...
```

Horizontal ellipses
These indicate that parts of the same turn are omitted:

```
...        D:    ...as I say yu-yuh go down steep...
```

Vertical ellipses
These indicate that intervening turns have been omitted:

```
.          Mum:   ↑oh: dea:r. Thursday.
.
.          (20 lines omitted))

           Mum:   Oh: they make you pay f'r putting it on again tog..
```

Numbering of lines or turns
This is done for convenience of reference. Intervals within talk are also numbered.

```
44     Les:   Yes well we sent the money straight away
45            (0.4)
46     Les:   ·p
```

Bibliography

I. Sources of data

BNC (British National Corpus). https://www.english-corpora.org/bnc/

CLAP (Corpus de LAngues Parlées en interaction). http://clapi.ish-lyon.cnrs.fr/

COCA (Corpus of Contemporary American English). https://www.english-corpora.org/coca/

Reverso context. https://context.reverso.net/traduction/

II. Articles and monographs

Aaltio, Maiji-Hellikki. 1987. *Finnish for foreigners I*. Helsinki: Otava.

Akatsuka, Noriko. 1986. Conditionals and epistemic scale. *Language* 61(3): 625-639.

Alvarez-Caccamo, Celso. 1996. The power of reflexive language(s): Code displacement in reported speech. *Journal of pragmatics* 25: 33-59.

Andrason, Alexander. 2017. The "exotic" nature of ideophones: from Khoekhoe to Xhosa. *Stellenbosch Papers in Linguistics* Vol. 48: 139-150.

Argoud, Line. 2010. Réalité des idéophones anglais (phonesthèmes) : propositions dans le cadre d'une approche de linguistique cognitive. In Jacqueline Perceboix, Isabelle Richard et Monique De Mattia-Vivès (eds.), *La Production et l'analyse des discours*. E-rea (revue électronique d'études sur le monde anglophone).

Bakhtin, Mikhaïl. 1981. *The Dialogic Imagination: Four essays.* Austin and London: University of Texas Press.

Bakhtin, Mikhaïl. 1986. *Speech Genres and Other Late Essays*. Austin, Texas: University of Texas Press.

Banfield, A. 1982. *Unspeakable Sentences: Narration and Representation in the language of fiction*. Boston, MA: Routledge and Kegan Paul.

Bazzanella, Carla. 2011. Redundancy, repetition, and intensity in discourse. *Language Sciences* 33: 243-254.

Bennett-Kastor, Tina L. 1994. Repetition in Language Development: From Interaction to Cohesion. In Barbara Johnstone (ed.), *Repetition in discourse,* volume 1: 155-171. Norwood: Alex Publishing Corporation.

Bergman, Brita & Dahl, Östen. 1994. Ideophones in Sign language? The place of reduplication in the tense-aspect system of Swedish Sign language. In Carl Bache, Hans Basbøll and Carl-Erik Lindberg (eds.), *Tense, Aspect, and Action*: 397-422. Berlin: De Gruyter Mouton.

Blagrove, Mark & al. 2019. Testing the Empathy Theory of Dreaming: The Relationships Between Dream Sharing and Trait and State Empathy. *Frontiers in psychology* 10 (article 1351.)

Bottineau, Didier 2013. Pour une approche énactive de la parole dans les langues. *Langages* 4 (192): 11-27.

Buchstaller, Isabelle. 2001. He goes and I'm like: The new quotatives re-visited. Paper presented at NWAVE 30, University of North Carolina.

Chafe, Wallace. 1994. *Discourse, Consciousness, and Time*. Chicago: University of Chicago Press.

Clift, Rebecca. 2007. Getting there first: non-narrative reported speech in interaction. In Elizabeth Holt and Rebecca Clift (eds.), *Reporting Talk*: 120-149. New York: Cambridge University Press.

Cohen-Achdut, Miri. 2019. Self-quotations and politeness: The construction of discourse events and its pragmatic implications. *Text&Talk* 39(3): 341-362.

Comrie, Bernard. 1996. Conditionals: A Typology. Elizabeth Traugott et al. (eds.). *On Conditionals*. Cambridge University Press.

Coppen, Peter-Arno & Foolen, Ad. 2012. Dutch quotative *van*: Past and Present. In Isabelle Buchstaller and Ingrid van Alphen (eds.), *Quotatives: Cross-linguistic and cross-disciplinary perspectives*. Amsterdam: John Benjamins. 259-280.

Couper-Kuhlen, Elizabeth. 2007. Assessing and accounting. In Elizabeth Holt and Rebecca Clift, (eds.), *Reporting Talk*: 81-119. New York: Cambridge University Press.

Dansieh, Solomon A., 2011. The pragmatics of reduplication: Implications for translating. *Journal of pragmatics*: 164-174.

Dingemanse, Mark. 2015. Ideophones and reduplication: Depiction, description, and the interpretation of repeated talk in discourse. *Studies in language* 39(4). 946-970.

Di Paolo, Ezequiel A., Clare Cuffari, Clare & De Jaegher Hanne. 2018. *Linguistic Bodies: the continuity between Life and Language.* Cambridge: The MIT Press.

Fischer, Olga. 2011. Cognitive iconic grounding of reduplication in language. In Pascal Michelucci, Olga Fischer and Christina Ljungberg (eds.), *Semblance and signification*, 55-82, Amsterdam: John Benjamins.

Fonagy, Ivan, 1986. Reported speech in French and Hungarian. In F. Coulmas (ed.), *Direct and indirect speech*: 256-309. Berlin/New York/Amsterdam: Mouton de Gruyter.

Fuchs, Thomas. 2017. The "as-if" Function and its loss in Schizophrenia. In Michela Summa, Thomas Fuchs, Luca Vanzago (eds.), *Imagination and social perspectives*: 83-99. New York: Routledge.

Galatolo, Renata. 2007. Active voicing in court. In Elizabeth Holt and Rebecca Clift (eds.), *Reporting Talk*: 81-119. New York: Cambridge University Press.

Gallagher, Shaun. 1986. Body image and Body schema: a conceptual clarification. *The journal of mind and behavior* 7(4): 541-554.

Garnier, Catherine & Mori, Toshiko. 2007. *Japanese with ease: volume one*. Chennevières-sur-Marne Cedex: Assimil.

Goffmann, Erving. 1981. *Forms of talk*. Oxford: Blackwell.

Gratier, Maya. 2009. Du rythme expressif à la narrativité dans l'échange vocal mère-bébé. *Champ psychosomatique* 54: 35-46.

Guerra, Victor. 2018. *Rythme et intersubjectivité chez le bébé*. Toulouse: éditions Érès.

Güldemann, Thomas. 2001. Quotative Constructions in Afican Languages: a Synchronic and Diachronic Survey. Unpublished 'Habilitationsschrift'. University of Leipzig.

Haakana, Markku. 2007. Reported thought in complaint stories. In Elizabeth Holt and Rebecca Clift (eds.), *Reporting Talk:* 150-177. New York: Cambridge University Press.

Histoire qui sont maintenant du passé. 1968. Traduction Bernard Frank. Paris: Gallimard.

Holt, Elizabeth, 2007. 'I'm eyeing your chop up mind': reporting and enacting. In Elizabeth Holt and Rebecca Clift (eds.), *Reporting Talk*: 47-80. New York: Cambridge University Press.

Hopper, Robert & Glenn, Philipp J. 1994. Repetition and Play in Conversation. In Barbara Johnstone (ed.). *Repetition in discourse*, volume II: 29-40. Norwood: Alex Publishing Corporation.

Jakobson, Roman. 1960. Concluding Statement: Linguistics and Poetics. In Thomas A. Sebeok (ed.), *Style in Language*: 350-377. Cambridge, MA: MIT Press.

Janowitz, Naomi. 1993. Re-creating: the metapragmatics of reported speech. In John A. Lucy (ed.), *Reflexive language*, Cambridge: Cambridge University Press.

Kauffmann, Charles A. 2015. Reduplication reflects uniqueness and innovation in language, thought and culture. Encyclopedia of writing systems and Languages. Publisher Simon Ager Omniglot – online. https://omniglot.com/language/articles/reduplication.htm

Kim, Jin Hyun. 2020. From the body image to the body schema, from the proximal to the distal: embodies musical activity toward learning instrumental musical skills. Frontiers in psychology. https://doi.org/10.3389/fpsyg.2020.00101

Kita, Sotaro. 1993. Two-dimensional semantic analysis of Japanese mimetics. *Liguistics* 35: 379-415.

Knox, Laurie. 1993. Repetition and Relevance: Self-repetition as a strategy for initiating cooperation in nonnative/native speaker conversations. In Barbara Johnstone (ed.). *Repetition in discourse,* volume I: 195-206. Norwood: Alex Publishing Corporation.

Labov, William. 1972. *Language in the inner city.* Philadelphia: University of Pennsylvania Press.

Langer, S. 1953. *Feeling and Form: A Theory of Art*. New York: Charles Scribner's Sons.

Larrory-Wunder, Anne. 2016. Exclamation et intersubjectivité en allemand. *Revue de sémantique et pragmatique* 40: 79-98.

Lucy, John A. 1993. Metapragmatic presentationals: reporting speech with quotatives in Yucatec Maya. In John A. Lucy (ed.)., *Reflexive language*: 91-125. Cambridge: Cambridge University Press.

Lucy, John A. 1993. Reflexive language and the human discipline. In John A. Lucy (ed.)., *Reflexive language*: 9-32. Cambridge: Cambridge University Press.

Machi, Saeko. 2008. *How repetition operates in Japanese and English: Introducing different cultural orientations towards conversation*. International spring forum 2008.

Mathis, T. & Yule, G. 1994. Zero quotatives. *Discourse processes* 18: 63-76.

Maynard, Senko K. 1993. *Discourse Modality: subjectivity, emotion, and voice in the Japanese language*. Amsterdam: John Benjamins.

Maynard, Senko K. 1996. Multivoicedness in speech and thought representation: The case of self-quotation in Japanese. *Journal of pragmatics* 25: 207-226.

Maynard, Senko K. 2002. *Linguistic Emotivity: Centrality of place, the topic-comment and an ideology of pathos in Japanese discourse*. Amsterdam: John Benjamins.

Mertz, Elizabeth. 1993. Learning what to ask: metapragmatic factors and methodological reification. In John A. Lucy (ed.), *Reflexive language*: 159-174 Cambridge: Cambridge University Press.

Miyeon, Ahn. 2005. A phonetically based account of English reduplication. *SNU working papers in English language and Linguistics* 4: 73-93.

Montaut, Annie. 2009. Reduplication and Echo-Words in Hindi. In Rajendra Singh (ed.), *Annual Review of South Asian Languages and Linguistics*: 21-90. Berlin: Mouton de Gruyter.

Moore, Robert E. 1993. Performance form and the voices of characters in five versions of the Wasco Coyote circle. In John A. Lucy (ed.), *Reflexive language*: 213-240. Cambridge: Cambridge University Press.

Norrick, Neal R. 1994. Repetition as a Conversational Joking Strategy. In Barbara Johnstone (ed.), *Repetition in discourse*, volume II: 15-28. Norwood: Alex Publishing Corporation.

Parmentier, Richard J. 1993. The political function of reported speech: a Belauan example. In John A. Lucy (ed.), *Reflexive language*: 261-286. Cambridge: Cambridge University Press.

Penelaud, Olivier. 2010. Le paradigme de l'énaction aujourd'hui: apports et limites d'une théorie cognitive révolutionnaire. *PLASTIR* 1: 1-38.

Piantadosi, Steven. T., Tily, Harry & Gibson, Edward. 2012. The communicative function of ambiguity in language. *Cognition* 122: 280-291.

Savan, David. 1980. La séméiotique de Charles S. Peirce. *Languages* 58: 9-23.

Sperber, Dan & Wilson, Deirdre. 1986. *Relevance: Communication and Cognition*. Oxford: Blackwell's.

Sperber, Dan & Wilson, Deirdre. 1995. Postface to the second edition of Relevance: Communication and Cognition: 255-280. Oxford: Blackwell.

Stépanoff, Charles. 2019. *Voyager dans l'invisible: techniques chamaniques de l'imagination*. Paris: Les empêcheurs de penser en rond.

Tannen, Deborah. 1989. *Talking voices*. Cambridge: Cambridge University Press.

Thibault, Paul J. 2011. First-order Languaging Dynamics and second-order Language: The distributed language view. *Ecological Psychology* 23: 1-36.

Tolins, Jackson & E. Fox Tree, Jean. 2014. Addressee backchannels steer narrative development. *Journal of pragmatics* 70. 152-164.

Tomlin, Russell S. 1994. Repetition in second language acquisition. In Barbare Johnstone (ed.). *Repetition in discourse,* volume I: 172-194. Norwood: Alex Publishing Corporation.

Urbanczyk, Suzanne. 2001. *Patterns of reduplication in Lushootseed*. New York & London: Garland Publishing.

Van huyssteen, Gerhard B. 2004. Motivating the Composition of Afrikaans Reduplications: a Cognitive Grammar Analysis. In Günter Radden and Klaus-Uwe Panther (eds.), *Studies in Linguistic Motivation*: 269-292. Berlin: Mouton de Gruyter.

Varela, Francsico. 1996. The emergent self. In John Brockman (ed.), *The third culture: Beyond the scientific revolution*: 209–223. New York: Simon & Shuster.

Watine, Marie-Albane. 2014. Les âges de la réduplication. In V. Magri-Mourgues & A. Rabatel, (eds.), *Pragmatique de la répétition* (Semen 38): 55-73.

Wilson, Deirdre. 2000. Metarepresentation in linguistic communication. In D. Sperber (ed.), *Metarepresentations: A multidisciplinary perspective*: 230-258. Oxford: Oxford University Press.

Yao, Bo and Belin, Pascal & Scheepers, Christoph. Silent reading of direct versus indirect speech activates voice-selective areas in the auditory cortex. *Journal of Cognitive Neuroscience* 23 (10). 3146-3152.