Modes of conference interpreting
Simultaneous and consecutive

Magdalena Bartłomiejczyk and Katarzyna Stachowiak-Szymczak

Introduction

Conference interpreting is carried out in two modes which fundamentally differ in their timing of the source and target texts: consecutive and simultaneous. In consecutive, the original speaker either pauses to enable the interpreter to deliver the target text in fragments, or a short speech may be rendered as a whole after the speaker has finished. In simultaneous, the speaker does not stop for the interpretation, which is delivered concurrently.

Consecutive interpreting requires extra time. Just one target language means that the time of the meeting will have to be almost doubled, if everything is to be interpreted (the interpretation should take anywhere between 2/3 and 3/4 of the time taken by the original, cf. Jones 1998: 40). Simultaneous interpreting, by contrast, allows multiple target language versions to be provided by interpreters working from different booths, without prolonging the meeting. By the same token, cognitive processes vary to some extent in these two modes and they also differ in the sub-tasks they include, such as note-taking, which is usually associated with the consecutive mode and stereotypically absent in simultaneous interpreting.

That notwithstanding, these and other typical features occur in what is regarded as ‘classic’—i.e. the most prototypical—simultaneous and/or consecutive interpreting, i.e. simultaneous interpreting in a booth, and a several-minute-long consecutive interpreting with a notepad. In turn, both simultaneous and consecutive interpreting can be further broken down into several ‘sub-modes’, sometimes (depending on the classification) regarded as modes in their own right.

The great time saving afforded by the simultaneous mode has made it the dominant form of interpreting, to the extent that many conference interpreters hardly ever perform consecutive interpreting nowadays (in Neff 2014, freelancers reported that over 85 per cent of their assignments were in the simultaneous mode, while staff interpreters reported 96 per cent). However, interpreters are generally required to be able to work in both modes, and both are taught in interpreting programmes, as skills in consecutive are often treated as a prerequisite for starting to learn simultaneous (see, e.g. Gillies 2019; Setton & Dawrant 2016b).
Simultaneous interpreting

Simultaneous interpreting functions as the epitome of conference interpreting, although it is also sometimes used in community interpreting settings (e.g. courts, educational institutions). The term is typically used to refer to interpreting between spoken languages, but in fact sign-language interpreting is most frequently also simultaneous as the source and target texts are delivered at the same time (although in different modalities, see Turner, Grbić, Stone, Tester & de Wit, Chapter 38, in this volume). Historically, in conference interpreting, this mode developed later than consecutive interpreting (see Baigorri-Jalón, Fernández-Sánchez & Payás, Chapter 1, in this volume), but it plays a more important role today and “is associated with higher professional status, higher levels of pay, and opportunities to travel and to be part of high-profile events” (Setton & Dawrant 2016b: 244).

Traditional simultaneous interpreting requires at least one booth to which the sound from the floor is transmitted. The interpreter listens to the source text through headphones and speaks into a microphone, and their voice is transmitted to the listeners, who also use headphones. If there are multiple booths, the listeners can switch among different channels (i.e. interpretations into different languages). Apart from built-in booths placed in conference halls, there are also mobile ones that can be quickly installed as needed. These do not require wiring for sound transmission as the system is based on infrared technology. ISO standards exist for both stationary (ISO 2603:2016; ISO 2016a) and mobile (ISO 4043:2016; ISO 2016b) booths to ensure workplace comfort and interpreters’ well-being.

In what sense is it simultaneous?

The simultaneity in simultaneous interpreting is to be understood in a very specific way. The speaker and the interpreter produce their versions of the text concurrently, but corresponding chunks of information (propositions) are not provided precisely at the same moment. The time lag between them is called the ear-voice span (EVS) or décalage. It is not and should not be constant during an interpreting task and EVS adjustment constitutes an important skill to be mastered by aspiring interpreters. As outlined by Gile (2009: 204), prolonging the EVS enables the interpreter to receive more information and may therefore facilitate comprehension, it may also give them more time to plan the target language version. On the other hand, prolonged EVS produces more strain on the interpreter’s memory. Shortening the EVS is recommended for fragments that are particularly difficult to remember, e.g. enumerations of numbers (Jones 1998: 131).

Measurements of EVS were undertaken very early in the history of interpreting studies (e.g. Oléron & Nanpon 1965/2002) and show that most of the time, the interpreter stays 2–4 seconds behind the speaker, the maximum lag, however, exceeds 10 seconds (Timarová 2015: 419). EVS is influenced by multiple factors, including syntactic differences between the languages involved, the interpreter’s individual cognitive style and expertise, speech rate, etc. (see, e.g. Setton & Dawrant 2016a: 304).

Interestingly, sometimes the interpreter may provide a chunk of speech before the speaker. Anticipation (see, e.g. Chernov 2004; Kalina 1998) becomes possible either because the interpreter recognizes the beginning of a familiar linguistic structure (such as a proverb), or because they predict the content on the basis of their extralinguistic knowledge (resulting from preparation for the conference, previous experience with the same speaker, etc.) (see Hodzik & Williams, Chapter 26, in this volume).
Early researchers (e.g. Barik 1973; Goldman-Eisler 1967) also inquired about the overlap between the original and the interpretation, hypothesizing that the interpreter might mainly squeeze their output into the speaker’s pauses to avoid concurrent language production. This hypothesis was rejected, however, when Gerver (1975) showed that these pauses where in fact too short, and various other studies measured the overlap ranging from 64 to 90 per cent (Pöchhacker 2004: 116; Setton 1999: 27–28). Simultaneous listening and speaking are likely to impress the audience, but experimental research shows that this subskill is relatively easy to acquire through exercise. In a study by Kurz, complete beginners (interpreting students who passed a screening test to be admitted to the programme) needed just 4 months to nearly reach the level of experienced interpreters (1996: 124).

**Processing in the simultaneous mode**

Linguistic processing in the interpreter’s mind is not directly observable. Concepts and models related to cognitive processing may be more or less compatible with interpreters'/teachers'/trainees’ intuition about simultaneous or consecutive interpreting, but they are very difficult to prove empirically (see, e.g. Pöchhacker 2004: 107–108).

Seleskovitch (1968/1978) presents interpreting as very similar to everyday language production. According to her, the main difference is that the meaning to be expressed in words comes from an external source instead of from one’s own mind. Comprehension of the source text happens spontaneously; the words of the original speaker are transformed into ideas (‘deverbalization’), and these ideas are given a new verbal shape. Consequently, it seems a fairly easy task that could be performed by any individual with excellent knowledge of the source and target languages. This vision, however, will in most cases be dispelled when a fluent bilingual actually tries to perform simultaneous interpretation. The inherent difficulty is also attested to by the fact that even experienced professional interpreters sometimes make grave errors even if the source text does not seem to present comprehension problems (see Gile 2009: 157–158).

The difficulty of simultaneous interpreting lies in extreme multitasking. Humans have limited cognitive capacities. Tasks that we find easy when done in isolation (e.g. talking to a friend, driving a car, finding a route on the map) may prove impossible to combine, or at the very least our performance of these tasks will deteriorate, which may affect some or all of the concurrent tasks. This is because each task that is non-automated requires some of our mental resources, which are only available in a limited supply. This common-sense realization, confirmed by cognitive psychology, lies at the foundation of arguably the most popular processing model: Gile’s Effort Models (2009: 158–171).

Gile describes simultaneous interpreting as consisting of four components: (1) Listening and analysis Effort; (2) Production Effort; (3) Memory Effort; and (4) Coordination Effort. All of them co-occur most of the time, requiring non-automated, conscious processing and compete for the interpreter’s processing capacity, which is limited (see Ricarditi, Chapter 27, in this volume). At any given moment, the processing capacity has to be allocated so as to cover the individual requirements for each Effort. Momentary peaks in the individual requirements are therefore likely to lead to saturation (the total processing capacity is insufficient) or to individual deficits (the processing capacity is not allocated correctly). The latter may occur, for example, when the interpreter strives to produce a very elegant target language solution and fails to listen attentively enough, being preoccupied with production. Consequently, they will more likely than not miss the incoming fragment of the source text.
Numerous other processing models exist. The earliest ones (Gerver 1975; Moser-Mercer 1978) endeavoured to present simultaneous interpreting as a decision-making process in the form of workflow charts, attributing much importance to the interaction of top-down and bottom-up processing and to monitoring. More recent, widely acknowledged models include the cognitive-pragmatic one by Setton (1999), drawing on Relevance Theory, cognitive semantics, mental models theory and speech-act theory. This model was meant to reconcile “the dimension of cognitive coordination and the transmission of meaning in SI though a better understanding of intermediate representation” (Setton 1999: 4). Though the model is sequential, all the processes are conceptualized as largely overlapping, with situational and world knowledge playing a role all the time. Among the newest models, Seeber’s Cognitive Load Model (2011) has generated much interest. It questions Gile’s (2009) assumption that “all tasks involved in the SI process draw on and the same pool of undifferentiated resources” (Seeber 2011: 189). The basis of this model is a “conflict matrix” mapping tasks that interfere with one another at a given stage of the interpreting process, taking into consideration the modality of each task (e.g. auditory, cognitive, verbal). Seeber proposes quantifying the cognitive demands for the tasks and calculates “conflict coefficients”, which are the highest for tasks of the same modality.

Constraints of the mode, problem triggers and how interpreters may overcome them

Apart from the interpreter’s cognitive constraints mentioned earlier, the linearity constraint and the time constraint are of crucial importance. Linearity (see Gumul 2017: 125–132; Jones 1998: 72) means that the interpreter only receives subsequent chunks of information as the text develops. They do not know where the speaker is heading, but depend on making informed guesses (anticipation is operating at all, including more global levels); (see also Hodzik & Williams, Chapter 26, in this volume). As the audience will be suspicious of longer pauses, some target text very often has to be produced before the interpreter comprehends the speaker’s intended meaning—this is when “padding” (Setton 1999: 50), for example, adding possibly neutral elements, such as polite forms of address, may be helpful.

The time constraint relates to the pace imposed by the original speaker, and is closely related to their speed of delivery (and information density, which may sometimes be high even though the speech rate is moderate, as in enumerations). It has been experimentally determined that the comfortable speech rate for an English source text lies between 95 and 120 words per minute (Pöchhacker 2004: 129–130; see also Barghout et al. 2015). Many speakers speak much faster than that, particularly if given the floor for a limited time. During the plenary debates of the European Parliament, for example, the average speech rate is about 150 words per minute (Monti et al. 2005). In most simultaneous booth systems, interpreters are able to inform the speaker and/or the chair of the meeting that they are struggling to keep up by switching on a signal light, but typically the speaker will only slow down for a short time (or even ignore the request).

Problem triggers (Gile 2009: 171) are those elements that are likely to cause difficulties to interpreters. They may relate to the speech as a whole (e.g. strong foreign or regional—that is, non-standard—accent, poor logic, fast speed) or occur locally (e.g. numbers, names, enumerations). The difficulty results either from increased processing capacity requirements or from signal vulnerability (e.g. unfamiliar names and numbers will be very difficult to reconstruct from the context if missed due to a momentary lapse of attention).

Interpreters develop methods to overcome such recurrent problems, referred to in the literature as ‘strategies’ (Bartłomiejczyk 2006; Kalina 1998; Riccardi 2005), ‘coping tactics’
(Gile 2009) or ‘techniques’ (Jones 1998). These may be applicable to very specific problem triggers—for instance, many interpreters note down numbers to facilitate their processing (Gile 2009), or work more globally—such as in the case of ‘visualization’ (Bartłomiejczyk 2006; Seleskovitch 1968/1978), which consists of forming imaginations related to the content of the source text. Some strategies typical for simultaneous interpreting have already been mentioned here, such as EVS adjustment, anticipation and padding. We will briefly discuss some more examples that reappear in various classifications (for more detailed information, see Riccardi, Chapter 27, in this volume).

Strategies that aid comprehension include “inferencing” (e.g. Gile 2009; Kalina 1998), which involves “trying to reconstruct the fragments of the original message which were not heard, not understood or forgotten by the interpreter on the basis of the context … or world knowledge relevant to the topic of the speech” (Bartłomiejczyk 2006: 160). A more extreme approach to the same type of problem is “parallel reformulation” (Bartłomiejczyk 2006; Gile 2009), in which the interpreter replaces the lost fragment with their own improvisation on the topic, which possibly has to be kept non-committal.

Strategies related to production include “transcodage” (Bartłomiejczyk 2006; Kalina 1998)—closely following the structure of the original message, “generalization” (Gile 2009; Kalina 1998)—replacement of an item with a superordinate term (e.g. an exact number such as ‘321’ might be rendered as ‘over 300’), and “paraphrase” (Bartłomiejczyk 2006; Gile 2009)—providing a longer explanation of a word whose equivalent the interpreter is not able to retrieve.

“Omission” is an “emergency strategy” (Kalina 1998) present in practically every classification. When applied strategically, omission is not haphazard, but instead involves selective deletion of information that is redundant (being repetitive or available elsewhere, for example, in the slides accompanying a presentation) or is relatively less important (Korpal 2012).

Cooperation of boothmates and larger interpreting teams

The level of attention required for simultaneous interpreting cannot be maintained for long, so interpreters work in pairs or threes and change every 20–30 minutes (or more often, if the source texts are particularly challenging). Moser-Mercer et al. (1998) found that interpreting turns longer than 30 minutes resulted in increased stress levels and deteriorating output quality. Accepting a simultaneous assignment during which the interpreter will have to work alone is considered highly unprofessional: AIIC (2015), in fact, recommends teamwork for both simultaneous and consecutive interpreting, specifying that one interpreter should be recruited only “under exceptional circumstances and provided the principles of quality and health are taken into full consideration”.

Interpreters sharing the same booth typically work into the same target language, but not necessarily from the same source language(s). In meetings involving numerous working languages, sometimes three or even four interpreters work in the same booth. In such cases, dividing the workload among boothmates is a very complex issue (see Duflou 2016). Interpreters who are not ‘on the mike’ often stay in the booth and they may offer some help to the colleague who is interpreting, for example, by noting down numbers and/or names in an informationally dense speech, or quickly checking target language terminology online. The COVID-19 pandemic has considerably disrupted interpreters’ teamwork routines, as due to the risk of infection resulting from sharing the same enclosed space, they are now often assigned individual booths (Hecht 2020) and, at best, see each other through a glass wall between neighbouring booths. Teamwork may be even more difficult to coordinate in remote interpreting, which has
also become immensely popular during the pandemic (see Seeber & Fox, Chapter 35, in this volume).

In international organizations, such as the EU, interpreting services are divided into language units likely to function as ‘communities of practice’ (i.e. groups whose members engage in the same activity directed at the same goal, while sharing “stories, histories, discourses, concepts, tools, styles, ways of doing things” Duflou 2016: 16). Interpreters from the same unit repeatedly cooperate with the same colleagues, which leads to the emergence of common norms and even some formulaic interpreting solutions (e.g. Henriksen 2007). There is also contact and cooperation between different units, especially if relay and retour interpreting (see below) are regularly used.

**Relay interpreting**

If a meeting has two or more booths, it becomes possible to interpret between two languages indirectly, using a third language as an intermediary. When there is a need for an interpretation, for example, from Polish into Hungarian, and no interpreter able to do this is available, the task can be entrusted to two interpreters. One will interpret from Polish into English, German, or French (the most popular ‘pivot languages’ in Europe), and the other will use this interpretation as a source text to interpret into Hungarian.

This method has some obvious disadvantages, due to which its use is discouraged (AIIC 2018b). These include evidently longer time-lag and “the added complexity resulting from the doubling of an already complex process”, which might render relay “doubly prone to errors or losses of message integrity” (Čeňková 2015: 340; see also Song & Cheung 2019 for an experimental study comparing regular and relay simultaneous interpretations). Nevertheless, it becomes a sheer necessity in meetings with very large language regimes, such as plenary sessions of the European Parliament held in 24 working languages. During these debates each booth is manned by three interpreters, and everything said on the floor has to be transferred into 23 languages. This means that each of the three interpreters, working into their mother tongue, would have to cover seven or eight source languages, which, moreover, would not overlap with those of their boothmates. Such an arrangement is simply unfeasible. Still, as outlined by Graves (2013), the EU institutions want to limit the use of relay as much as possible, and therefore, encourage their interpreters to acquire multiple passive languages (on average, they have four) (see Graves, Pascual Olaguíbel & Pearson, Chapter 8, in this volume). Similarly, AIIC discourages the use of relay unless direct interpreting is impossible (AIIC 2018b).

Acting as a ‘pivot’ for relay is particularly stressful (Čeňková 2015: 341), as the interpreter realizes that colleagues from other booths depend on their output and might be negatively affected by low quality. At the same time, relay is typically done from less popular languages, which means the interpreter will often be performing retour (see below).

**Consecutive interpreting**

Consecutive interpreting involves subsequent source and target text production (by speaker and interpreter), albeit concurrent bilingual and multimodal processing (see below). Consecutive interpreting is conducted in combinations of both spoken and signed languages. Simultaneous interpreting is the mode of choice during multilingual conferences (Russell & Takeda 2014: 103), while organizers of a bilingual event may consider consecutive interpreting as a good option. Gillies (2019: 7) enumerates several reasons why consecutive might be preferred over simultaneous. These include lower costs, better interaction between the interpreter
and the participants, and giving delegates who master both the languages some extra time to think before they have to respond. Consecutive interpreting may also be conducted without any technical aid, while during some events source text sound may need to be delivered via headphone(s) to the interpreter, and/or the interpreter may have to use a microphone (these should be a floor, desktop or lapel one, to guarantee maximum comfort and enable free hand movements; AIIC 2016).

What is consecutive in consecutive interpreting?

The ‘consecutiveness’ of this mode is best illustrated from the perspective of the target audience, who waits for the interpretation until the speaker finishes the speech or a fragment thereof. In other words, there is no overlap of speech streams in consecutive interpreting. Although source texts can be as short as a single word or phrase (Pöchhacker 2004: 18), consecutive conference interpreting usually involves longer chunks, frequently of several minutes. There is high variation in length, depending, inter alia, on what users are accustomed to on a specific market. For instance, chunks rendered at meetings requiring high participant involvement, such as workshops or business meetings usually last no more than several sentences. On the other hand, consecutive interpreters at international galas, lectures or interpreting politicians’ communications may be faced with several-minutes-long speeches. Having that in mind, some master’s programmes end with a post-training assessment conducted by means of speeches that last around 5 minutes, including the European Masters of Conference Interpreting (EMCI) Consortium where final exam texts for the consecutive mode last 5 to 7 minutes (EMCI 2017).

The source text to be rendered can be perceived as a certain “set” of ideas, logical connections, context and other elements which have to be reproduced in the target speech. In consecutive interpreting, elements of the source text can be processed, reorganized and structured into the target text relatively long after they are perceived by the interpreter who, contrary to the simultaneous mode, is aware of the full or larger parts of the text when doing the interpretation. On one hand, this allows the interpreter to avoid rendering redundant information, repetitions, etc., as well as enables reformulation and often reorganization of ideas for the sake of clarity (Albl-Mikasa 2008; Seleskovitch 1975). On the other hand, consecutive interpreting puts a strain on the interpreter’s memory (see below).

Language use, by contrast, is not consecutive in consecutive interpreting. Although target text production follows source text reception, interpreters may start encoding the target text before the source one ends (Ribas 2012). As a consequence, some fragments of the target text can be reformulated and reorganized several times prior to articulation. This is partially due to the phenomenon of non-selective language access, i.e. the activation of one language despite perception and/or production in another one (de Groot et al. 2000). To put it succinctly, despite two distinct languages being used for articulation in the source and target text production, neither of them is suppressed in the interpreter’s mind at any stage of the interpreting process. At the same time, while both languages are active, they differ as to the level of activation of their input and output mechanism (Grosjean 1997: 174). Such non-selective language access is manifested in bilingual notation. In other words, interpreters frequently make notes in two (or even more) of their working languages, which does not always correspond to the language articulated at a particular time (Dam 2004: 14ff.).

Multitasking that seems so characteristic of simultaneous interpreting is also prominently present here, as several activities are conducted concurrently. These involve, though are not always limited to: text decoding and encoding, memory operations, note-taking, note reading, keeping eye contact, gesturing, self-monitoring, etc. Some of these may call for extra cognitive
effort, while the co-existence of other operations may facilitate interpreting, which is explained in the next section.

**Processing in the consecutive mode**

Consecutive interpreting is frequently intuitively broken down into two main stages, corresponding to the part when the interpreter is listening and not speaking, and the part when the interpreter speaks to the audience. This division is represented in early approaches and models of, for instance, Alexieva (1998) who proposes phase 1 (listening and analysis, later renamed as comprehension) and 2 (target text production, later renamed as reformulation). Gile (2009: 175) further subdivides consecutive interpreting into listening and analysis, as well as target text production. In his Effort Models, listening and analysis overlap with note-taking and are accompanied by short-term memory operations. They lead to target production facilitated by note reading and supported by long-term memory operations. At the same time, it seems vital to remember that Gile’s Effort Models were initially designed for training, i.e. to facilitate trainees’ understanding, analysis and enhancement of their performance, not as a scientific tool (Gile 2018). Therefore, the model constitutes a conceptual illustration of consecutive interpreting, not a psycholinguistic or neurolinguistic one.

Consecutive interpreting is also a multimodal process, where language processing is embodied in activities, such as hand gestures, that are synchronized with language comprehension during notation and/or language production (see e.g. Pouw et al. 2014 for embodied language processing). Note-taking (see Ahrens & Orlando, Chapter 3, this volume), which supports the interpreter’s memory (see, e.g. Chen 2016), involves motor movements, while the notes themselves constitute visual input, which also makes consecutive interpreting multimodal. Furthermore, the source and target texts are received and produced in the auditory modality when two spoken languages are used, while the visual and motor modalities are added when the language combination involves a sign language. At the same time, interpreters make use of the visual modality in consecutive interpreting when reading not just their notes but also other visual material in the form of vocabulary lists, documents (usually in short consecutive mode or liaison) or slides prepared by the speaker, all of which may complement the interpreter’s notes. In addition to this, the source text speaker’s body language serves as a powerful source of input to the interpreter, as speech-accompanying gestures and facial expressions can facilitate language comprehension and provide for additional information (Alexieva 1998: 182; McNeill 1992). Finally, it has been found that consecutive interpreters resort to content visualization and mental imagery operations to facilitate information storage and recall (Jones 1998: 34–35; Kriston 2012).

**Constraints in the consecutive mode**

Consecutive interpreting is characterized by a relatively long period of time that passes between the moment a piece of information is received by the interpreter and the moment it is articulated in the target language. Therefore, information needs to be held in the interpreter’s memory, which puts a strain on the memory (Chuang 2008; Lambert 2004) despite the facilitative and supportive role of note-taking and mnemonics (Kriston 2012). Consequently, text duration adds to the general challenge of consecutive interpreting (Ribas 2012: 820). Despite its support function, notation also requires attention and so cognitive capacity likewise has to be assigned to note-taking and note-reading, as well as to coordination of the two. In fact, coordination of all the various simultaneous bilingual processes constitutes a major challenge in the
consecutive mode. Apart from this, elements that are challenging in consecutive interpreting are not dissimilar from those that constitute a challenge in the simultaneous mode: In both modes, they are commonly known as problem triggers (see above and Gile 2009).

Unlike in simultaneous interpreting, both the source and the target text are accessible to the audience, so that errors or omissions are likely to be noticed by listeners who have some command of both the languages. Second, problem triggers may in general increase the amount of processing capacity dedicated to either comprehension or note-taking and consequently impede the memorization of the subsequent source text elements (Gile 2009).

Lastly, some difficulties may arise due to the specificity of the assignment. An inexhaustive list of events with high degree of unexpectedness includes live shows, dynamic visual presentations, animal trainings, and sporting events, etc. Other problems can include, but are not limited to, issues with auditory perception, or the stressogenic effect of interpreter exposure to a large audience. In fact, interpreter visibility, especially during on site (as opposed to online) events requires efficient monitoring of body language and personal space use; this may be a stressor in the consecutive mode as opposed to the simultaneous mode where the interpreter remains ‘hidden’ in the booth. At the same time, many of these remain solvable by training and efficient client-interpreter cooperation.

On the variety of types, modes and combinations

Working languages (A, B, C) and the controversial status of A–B interpreting

The interpreter’s working languages are typically classified as outlined by AIIC, the International Association of Conference Interpreters (AIIC 2018a). A is the L1 or native language. Symmetrical bilinguals who have acquired two languages as children and have continued using both through their education and work life are considered to have two As, but this is relatively rare. B is a foreign or second language (L2) of which the interpreter has a near-native command (flawless grammar, mastery of a wide array of domains and little or no accent), so that they are able to interpret both from and into this language. C, on the other hand, is a foreign language (which can be a second or a third or more language) that the interpreter understands perfectly, but speaks less fluently. Therefore, they only interpret from C and not into C. Some conference interpreters have as many as 8 or 9 C languages (Neff 2014), a few of which they have learned when already working as interpreters. A and B are often referred to as active languages, and C as passive languages. This classification of working languages applies to both simultaneous and consecutive interpreting, but sometimes the same foreign language might be the interpreter’s B for consecutive and C for simultaneous, or, less typically, the other way round.

Many international organizations (e.g. the UN, the EU) favour interpreting into A, but making it a general rule proves unfeasible. For instance, when Finland joined the EU in 1995, it was not possible to find enough interpreters having Finnish as a C language, consequently, interpretations from Finnish into other languages had to be provided mostly by native Finnish interpreters. After the ‘big bang’ enlargement in 2004 and addition of official languages such as Hungarian and Maltese, into-B interpreting became even more common. On the private market, interpreters regularly work into B, while C languages are generally not in particularly high demand.

The belief in the superiority of B–A interpreting over A–B (also referred to as retour) used to be rather widespread and particularly pertinent for simultaneous, due to the fact that in other modes interpreters are typically required to work in both directions as a matter of course.
Some Western European scholars such as Seleskovitch (1968; 1978) claimed that the quality of *retour* is so inferior in the simultaneous mode that it should neither be performed nor taught. In contrast, scholars from Eastern Europe (e.g. Denissenko 1989) argued that into-B interpreting was both easier and better due to the considerable comprehension advantage for the native language. Over time, the so-called directionality debate has become less heated in the sense that both B–A and A–B interpreting are now widely recognized as capable of ensuring adequate quality and fulfilling a genuine market demand. It is often believed that the status of the source and target languages as the interpreter’s A or B should not be considered independently of a number of other factors, such as the languages themselves and their (dis)similarities or the type of event to be interpreted (see, e.g. Kalina 2005). A good deal of research has been published on directionality effects for various language combinations (e.g. Bartłomiejczyk 2006; Godijns & Hinderdael 2005; Monti et al. 2005; Yinyin & Posen 2018), pointing to a number of differences in strategic processing that should be taken into account both in professional practice and in training.

**Traditional simultaneous and consecutive vs. other sub-modes**

‘Classic’, on site simultaneous interpreting (see Seeber & Fox, Chapter 35, in this volume) involves the use of an interpreting system with booths. In contrast, chuchotage or ‘whispered interpreting’ (see, e.g. Pöchhacker 2004: 19; Setton & Dawrant 2016a: 18–19) is the oldest type of simultaneous interpreting that does not require any equipment. The interpreter sits next to the delegate who needs interpretation and speaks in a low voice. Chuchotage has many disadvantages: it is only suitable for very small groups of listeners and creates some disturbance for those who listen to the source text. The interpreter may also be disturbed by noise in the room. The disadvantages can be reduced by using the bidule system (a wireless microphone for the interpreter connected with receivers and headphones for the audience), which enables the interpreter to sit further apart and to cater for the needs of numerous listeners. Still, it remains “ersatz SI without booths … for obvious reasons of inadequate sound quality and acoustic isolation” (Setton & Dawrant 2016a: 19).

Sight translation (sight interpreting being a more accurate, but less widespread term) involves rendering orally into another language a written text that, typically, the interpreter has not read before. Therefore, the simultaneity is not to be understood as concurrent delivery of the source and target text, but as concurrent reception and production by the interpreter. In terms of processing, Listening Effort is replaced by Reading Effort (Gile 2009: 179). The speed is not imposed by the speaker, but in conference settings the audience will still expect a fluent rendition in which the interpreter does not make pauses for reading the source text or planning the target text. Apart from its occasional use in conference settings, sight translation is more frequently employed in community interpreting, e.g. in courts.

Simultaneous interpreting with text may be indistinguishable from the traditional variety for the audience, but it is considerably different for interpreters. On the one hand, they do not have to rely exclusively on ephemeral auditory input and the strain on their memory is also smaller. Unless they receive the text at the very last moment, their preparation may be much more targeted, which boosts self-confidence. On the other hand, when looking at the text during interpreting they have to allocate their processing capacity to yet another channel: the visual one. In Gile’s terms, both Listening Effort and Reading Effort are thus active (2009: 181). In spite of trainers’ warnings that “the priority must be on listening” (Setton & Dawrant 2016a: 299) and that the written text, once read, should only serve as verification and, possibly, retrieval of items such as proper names and numbers, research suggests that interpreters focus
more on the visual than the auditory input (Chmiel et al. 2020). This entails certain risks: the speaker may depart from the submitted text, for instance, by skipping some parts or adding ad-libbed remarks, or the interpreter may fall very far behind the speaker while sight-translating.

Simultaneous consecutive interpreting (SimConsec), sometimes also called consecutive simultaneous, is a relatively new and rarely used hybrid sub-mode (see Pöchhacker 2015). It consists of using a recording device, such as a digital pen (Orlando 2014), to record the source text, and subsequently replaying it to the interpreter via a headphone (so that they do not have to take notes while the source text is produced). The interpretation is delivered after the original, so for the audience it resembles long consecutive interpreting. In terms of cognitive processing, though, it is closer to simultaneous, with the advantage of having heard the whole text earlier. SimConsec may be particularly useful for visually impaired interpreters as a replacement for long consecutive interpreting with notes (see Fantinuoli, Chapter 36, in this volume).

Confusingly, the same term, i.e. simultaneous consecutive, may also refer to the situation in which interpreting booths are used to provide several concurrent consecutive renditions in different languages—this is how the newly invented equipment was employed in its early days (Pöchhacker 2004: 18).

Liaison interpreting is frequently perceived as a sub-mode of consecutive. Liaison (usually) proceeds in the form of interpreting short chunks subsequently to the source text speaker, thus consecutively. It is usually conducted for a limited number of interlocutors, e.g. during business meetings, and characterized by frequent turn-taking and dynamics. It may be accompanied by or used in exchange with chuchotage (Pöchhacker 2013: 155). Examples of using the ‘mixed’ chuchotage-liaison mode include, e.g. panel discussions or workshops, when one direction serves a very small number of listeners (e.g. one or two), whereas the other direction serves a larger audience, and there is no equipment enabling chuchotage to be delivered via two channels to all listeners. In this case, chuchotage is conducted for the minority audience (mainly for the sake of meeting dynamism), while anything they say is interpreted in the consecutive mode for the larger audience. During liaison interpreting (particularly in courts), the interpreter may sometimes start producing the target text before the speaker has finished producing a chunk, which results in a partial overlap, half-way between a consecutive and a simultaneous rendition.

Finally, it seems vital to comment on possible language combinations and directionalities in consecutive interpreting. Consecutive can occur between spoken and signed languages. In short, interlocutors may receive and/or produce text in auditory modality (which boils down to spoken language), or in visual modality, i.e. sign language (see Turner, Grbić, Stone, Tester & de Wit, Chapter 38, in this volume) or text. Speech-to-text (STT) interpreting proceeds simultaneously from spoken language to written text. Its applications include, but are not limited to, a situation when a Deaf person uses spoken language for production and either does not use sign language or interacts with a person with no sign language competences. STT is not dissimilar to respeaking, i.e. rendering inter- or intralingual subtitles (sometimes in the form of longer chunks) from speech, based on speech recognition and the post-editing of subtitles (see Norberg et al. 2015; Romero-Fresco 2011).

Conclusion

Conference interpreting is primarily divided into two main, albeit not the only, existing modes: simultaneous and consecutive interpreting. Although they look highly divergent from the audience’s perspective, these two modes are not dissimilar in terms of language processing, language access and working memory operations. They both involve bilingual and multimodal
processing, and rely on multitasking, while both may also involve reading and notation. At the same time, there are some differences between the modes which seem to refer to their technical and organizational characteristics, such as the use of equipment. The key difference lies in the amount of source and target text overlap, which is non-existent in consecutive. Working in the consecutive mode usually involves high interpreter visibility, in contrast to simultaneous, with interpreters ‘hidden’ in the booth or, when doing chuchotage, not entering into direct interaction (including eye contact) with the audience.

Each mode possesses its unique challenges. Decision-making has to proceed faster in simultaneous interpreting, and the risk of interference from the source language seems higher. On the other hand, consecutive interpreting is likely to involve more stage fright, especially when performing in front of a large audience. Moreover, access to support in the form of conference materials and self-prepared glossaries is heavily limited, not to mention the possible lack of assistance from colleagues.

In this chapter, we have mainly considered the process aspect of the two modes. However, when the end product is examined, consecutive and simultaneous interpretations also display some differences. Experimental studies involving interpreting the same texts in each mode are relatively rare: Gile (2001) assesses simultaneous interpretations as more accurate than consecutive ones but is reluctant to generalize on the basis of his results due to the modest scope of the study, while Gumul (2012) has established that consecutive interpretations are significantly more prone to shifts in cohesion. Elsewhere, in comparing corpora of authentic simultaneous and consecutive interpretations, Lv and Liang (2019) highlight that the latter are more simplified in terms of lexical parameters and information density in relation to the source texts. They construe this as indicative of relatively higher cognitive load for consecutive interpreting. Importantly, simultaneous interpretations are often affected by what Shlesinger (1994) calls “interpretational intonation”, characterized negatively by pauses within grammatical constituents, fluctuating speed of delivery, anomalous pitch and stress, whereas consecutive interpretations are likely to be more pleasant to the ear. In general, a combination of process-, product-, agent- and training-oriented approaches would likely show the full range of characteristics of simultaneous and consecutive interpreting.

Further reading


References


International Organization for Standardization 2016a. Simultaneous interpreting—Mobile booths—
International Organization for Standardization 2016b. Simultaneous interpreting—Permanent booths—
Kalina, Sylvia 1998. Strategische Prozesse beim Dolmetschen: Theoretische Grundlagen, empirische
Fallstudien, didaktische Konsequenzen. Tübingen: Gunter Narr.
Korpal, Pawel 2012. Omission in simultaneous interpreting as a deliberate act. In A. Pym & D. Orrego-
Kriston, Andrea 2012. The importance of memory training in interpretation. Professional Communication
and Translation Studies 5 (1–2), 79–86.
Kurz, Ingrid 1996. Simultandolmetschen als Gegenstand der interdisziplinären Forschung. Vienna:
WUV-Universitätsverlag.
Lambert, Sylvie 2004. Shared attention during sight translation, sight interpretation and simultaneous
Lv, Qianxi & Liang, Junying 2019. Is consecutive interpreting easier than simultaneous interpreting?: A
Chicago Press.
Monti, Cristina, Bendazzoli, Claudio, Sandrelli, Annalisa & Russo, Mariachiara 2005. Studying directionality in simultaneous interpreting through an electronic corpus: EPIC (European Parliament
Moser-Mercer, Barbara, Künzli, Alexander & Korac, Marina 1998. Prolonged turns in interpreting:
902/AIICWebzine_2014_Issue64_5_NEFF_AIIC_statistics_Summary_of_the_2012_report_EN.pdf
(visited 5 August 2020).
Norberg, Ulf, Stachl-Peier, Ursula & Tiittula, Liisa 2015. Speech-to-text interpreting in Finland, Sweden
and Austria. Translation & Interpreting 7 (3), 36–49.
Orlando, Mark 2014. A study on the amenability of digital pen technology in a hybrid mode of
Toward a more embedded/extended perspective on the cognitive function of gestures. Frontiers in
August 2021).
Ribas, Marta Arumi 2012. Problems and strategies in consecutive interpreting: A pilot study at two
different stages of interpreter training. Meta 57 (3), 812–835.
50 (2), 753–767.
Romero-Fresco, Pablo 2011. Subtitling through Speech Recognition: Respeaking. Manchester/
Kinderhook: St Jerome.


