Journal of Phonetics 108 (2025) 101377

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/Phonetics

Investigating interlanguages beyond categorical analyses: Prosodic marking of information status in Italian learners of German ☆

Simona Sbranna *, Aviad Albert, Martine Grice

IfL-Phonetik, University of Cologne, Herbert-Lewin Straße 6, 50931, Köln, Germany

ARTICLE INFO

Article history:
Received 7 August 2023
Received in revised form 11 November 2024
Accepted 17 November 2024
Available online 6 December 2024

Keywords:
L2 prosody
Prosodic marking of information status
Prosodic prominence
Continuous phonetic parameters
Periodic energy
Categorical analysis

ABSTRACT

Previous studies report that Italian learners of German transfer their L1 prosody to their L2 when marking information status prosodically within noun phrases (NPs). However, these studies were based on a categorical analysis of accentuation based on the presence or absence of pitch accents, which might not provide the full picture of interlanguages, in which category boundaries are flexible and dynamically evolving. We elicited two-word NPs in two different information status conditions – given-new (GN) and new-given (NG) – in L1 German, L1 Italian, and L2 German. We performed a periodic-energy-informed analysis to explore speakers' continuous modulation of F0 and prosodic strength and additionally discuss the results for the interlanguage in categorical terms. Learners prosodically mark information status by modulating the F0 contour on the first word similarly to their L1. However, learners reduce the prosodic strength of the second word in the noun phrase across the board, i.e. irrespective of information status. This pattern resembles German deaccentuation, and indicates that the learners are using a salient pattern but are not associating it with the appropriate pragmatic function. The current study revealed patterns for L1 Italian learners of L2 German which did not emerge in previous categorical analyses of the intonation of Italian learners of German.

1. Introduction

Since languages differ in how prosody is used to mark information status, this prosodic skill might present a challenge to second language (L2) learners. To date, there are few studies on the prosodic encoding of information status in interlanguages, despite the fact that this encoding has important consequences for successful communication. Such studies also have considerable potential for educational applications above and beyond the theoretical knowledge gap that such studies can fill. Studies on the prosodic encoding of information status in interlanguages mostly report on a transfer of phonetic and phonological cues from the L1 to the learners' L2. They differ in the methodological approach to learners' prosodic realisations, in some cases providing gradient analyses and in others categorical.

In this contribution, we aim at enriching the existing body of knowledge by unveiling the complexity of second language acquisition (SLA) processes. We use an exploratory approach that provides information about entire structures (rather than merely selected positions), broken down into pitch-related and strength-related metrics that we derive from the acoustic signal. Using this approach as a first approximation, we go beyond the reduction that discrete categorical symbols in formal models of intonation necessarily convey, to gain a window onto the fine details of prosodic aspects of the interlanguage (a language produced by learners that is a unique linguistic system different from both learner's native and target languages, but still presenting elements of both of them, as in Selinker, 1972) with more objective, open-ended and theory-neutral measures of the continuous signal (rather than derived symbols).

Specifically, in this paper we investigate the prosodic marking of information status (in our case, new or given with respect

☆ This article is part of a special issue entitled: Linguistic & Cognitive published in Journal of Phonetics.

* Corresponding author.
E-mail addresses: s.sbranna@uni-koeln.de, a.albert@uni-koeln.de, martine.grice@uni-koeln.de

https://doi.org/10.1016/j.wocn.2024.101377
0095-4470/

to the context) within noun phrases by Italian learners of German, following the footsteps of the pioneering studies on Italian and German as L1 and L2 in Swerts et al. (2002) and Avesani et al. (2013, 2015), which used a categorical approach. To offer new insight on this issue, we use an innovative acoustic analysis to explore the modulations of F0 and prosodic strength used by learners to prosodically distinguish information status in production (previously used effectively to analyse the learners' native language, Italian, in Sbranna et al., 2023).

In the following section (Section 2), we review the background on prosodic marking of information status, first in second language research, then in both Italian (native language of the learners) and German (target language of the learners), and we describe the comparative approach we used to investigate the intonation of Italian learners of German. These accounts are followed by a description of the production study (Section 3), and the results of the acoustic and statistical analyses (Section 4). Finally, we present a discussion in Section 5, followed by our conclusions in Section 6.

2. Background

2.1. Prosodic marking of information status in second language acquisition

The evidence collected so far about the prosodic marking of information status in L2 is mostly relative to learners of English as L2, and the effects of proficiency reported are inconsistent. In particular, learners' prosodic encoding of information status has been found to differ from English native speech, both phonetically and phonologically.

Phonetic differences involve the use of F0 peak alignment, F0 height and F0 movement. For example, Korean learners of English produced a delayed pitch peak on new information (Trofimovich & Baker, 2006), probably owing to the influence of L1 pitch alignment patterns. Mandarin Chinese learners of English, instead, presented less difference in pitch excursion across new and given items in their interlanguage compared to L1 English speakers (Wennerstrom, 1998). The same result is reported for Spanish learners of English, whose pitch range on focused constituents is narrower than in native productions, without clear differentiation from the adjacent syllables. Moreover, these learners were also found to produce a falling contour on both new and given elements, while native speakers differentiated the information status using a fall and a low rise respectively (Verdugo, 2003). A study on Malay learners of English showed that the phonetic details of L2 rises resembled those in the speakers' L1 (Gut & Pillai, 2014).

From a phonological point of view, there seems to be a common tendency for non-native speakers to overaccentuate, regardless of information status (Austrian learners in Grosser, James, & Leather, 1997; Spanish learners in Verdugo, 2003; learners with various L1 backgrounds in Gut, 2009; Malay learners in Gut & Pillai, 2014). In interpreting this result, it is important to keep in mind that the target language is always English, in which accentuation is used to mark information status, whereas the learners' native languages might have a looser relation between accentuation and information status, and might follow other criteria for distributing accents in the utterance, leading to the accentuation of given elements in the L2. This is the case for Spanish and Malay learners accentuating given elements in final position, where final position is taken to be the default, or unmarked case, i.e. "the pattern that is chosen when there is no compelling grammatical or contextual reason to choose some other" (Ladd, 2008, p. 223).

In a cross-linguistic study on Dutch and French (Rasier et al., 2010), it emerged that not all learners tended to overaccentuate, and that both Dutch learners of French and French learners of Dutch transferred their L1 features in prosodically encoding the information status of noun phrases. Specifically, Dutch learners tended to use the less common French "focus accent" (more similar to their own L1), in which only one element of the noun phrase is accented, and never used the more common French "bridge accent", in which both elements of the noun phrase are accented. French learners also applied their own L1 strategy to the L2 in not deaccenting contextually given information. The difference between learner groups is explained by the *Markedness Differential Hypothesis* (Eckman, 1977), according to which, marked processes and structures (such as selectively accentuating to communicate pragmatic contexts) are more difficult to learn than unmarked ones (such as structurally accentuating by position in the default condition). Dutch learners therefore have an advantage over French learners in this regard because the presence of a marked process in a given system usually entails the presence of the unmarked process in the same system as well, but not the other way around. Another study conducted on languages other than English reached similar results (Turco, Dimroth, & Braun, 2015). While in L1 Italian polarity contrasts were marked through a verum-focus accent (i.e. contrastive pitch accents on the finite verb, p.466) in a minority of cases, German learners of Italian produced more verum-focus accents and transferred their L1 phonetic implementation by often deaccenting post-focal constituents instead of using post-nuclear pitch accents. Transfer was also found for Dutch learners of Italian, who preferred lexical markers as in their L1 (Turco, et al. 2015). From these findings based on L2s other than English, it appears that the influence of the L1, rather than a universal tendency (Gut & Pillai, 2014; Gut, Pillai, & Don, 2013), better accounts for the consistent overaccentuation found in L2 studies on the prosodic encoding of information status.

Finally, results are contradictory about the effect of proficiency on learners' prosodic marking skills. Some studies show that transfer of L1 features tends to decrease as language proficiency increases. For example, in a study comparing speakers of L1 Zulu with their L2 English, it was found that beginners did not mark information status within noun phrases through accents for contrastive or corrective focus in accordance with in their L1, while advanced learners tended to do so (Swerts & Zerbian, 2010). Likewise, advanced Japanese learners of English were able to map given information to deaccentuation and contrastive information to an L + H* accent in the same way as native English speakers in both rating and production tasks, whereas less proficient L2 learners only mapped these in the rating task (Takeda, 2018). In contrast, some other studies do not find an improvement in intonational competence with higher proficiency (Chen & Bi, 2008; Verdugo 2003).

2.2. Prosodic marking of information status in Italian

The extensive evidence of a close relationship between prosodic accentuation and information status found in West Germanic languages (Halliday, 1967; Terken, 1984; Ladd 1996; Cruttenden, 2006; Fery & Samek-Lodovici 2006) has influenced successive research questions on Romance languages. A clear example is the use of the term "re-accentuation" of given elements in early research on Romance languages, to contrast with "deaccentuation" in West Germanic languages (Cruttenden, 1993; Swerts, Avesani, & Krahmer, 1999), implying that no accent is in some sense the subtraction of an accent that is assigned by default (even if Cruttenden 1993, 1997, himself observed that these consistent associations are in fact "preferred" patterns, allowing for alternatives). The widespread interest in the phenomenon of deaccentuation also included some studies on Italian, starting from Ladd's observation that Italian seems to allow deaccentuation of entire syntactic constituents, i.e. full clauses or noun phrases, but blocks deaccentuation within syntactic constituents (Ladd, 1996), even if Italian tends to preferentially use word order to mark information status (Cruttenden, 1993; Ladd, 1996, 2008: Face & D'Imperio, 2005).

Some acoustic–phonetic studies provided support for Ladd's observation, reporting a reduced F0 range and duration in post-focal falling contours in northern Italian varieties (Farnetani & Zmarich, 1997), or a low and flat F0 contour on post-focal given elements in Tuscan (Avesani et al. (1995); Hirschberg & Avesani, 1997) and Neapolitan Italian (D'Imperio & House, 1997), in both full clauses and single phrases.

Studies based on categorical phonological analyses, however, came to different conclusions. Investigations of accentuation and accent types in Tuscan (Avesani, 1997) and Roman Italian (Avesani & Vayra, 2005) found given elements to be accented irrespective of their function and position in the discourse, with some exceptions found in the latter variety, namely post-focal deaccentuation in sentences with fronted foci. Successive acoustic–phonetic investigations on Tuscan Italian showed that post-focal given elements occurring as a head of the prosodic domain (i.e., in metrically strong position) had an increase in duration, with more pronounced formants and spectral emphasis, but no evidence of F0 movement (Bocci, 2013; Bocci & Avesani, 2011, 2015). In an attempt to disentangle this complex picture, Bocci (2013) interprets the low and flat F0 contour found on post-focal given elements as an L* pitch accent, rather than as deaccentuation. A similar suggestion is made for Neapolitan Italian (D'Imperio, 1999), in which the flat post-focal region is interpreted as a compressed, downstepped version of the perceptually non-salient H + L* phrase accent.

The evidence collected across different varieties of Italian is inconsistent and sometimes contradictory due to different methodological approaches and the great deal of dialectological variation (Bertinetto & Loporcaro, 2005; Canepari, 1980; Giordano, 2006; Lepschy & Lepschy, 1977; Magno-Caldognetto, Ferrero, Lavagnoli, & Vagges, 1978; Savino, 2012; for a comprehensive review see Gili Fivela et al., 2015; Vietti, 2019).

2.3. Prosodic marking of information status in German

In German, despite the relative flexibility of its word order, prosody is the main linguistic marker of information status. However, the traditional strong association between givenness and accentuation assumed for West Germanic (Allerton, 1978; Cruttenden, 2006) has been relativised by several investigations (Baumann, 2005; Baumann & Hadelich, 2003; Féry & Kügler, 2008; Wagner, 1999), showing that even if deaccentuation is the most appropriate and common way to mark givenness, different options are available. For example, in a study on both read and spontaneous speech, Baumann & Riester (2013) found fewer deaccentuation cases than expected from the literature. They suggest a gradient scale of prosodic prominence, realised through a range of different accent types (including deaccentuation), mapping onto different degrees of activation of a referent (Baumann, 2006; Baumann & Riester, 2013). In turn, different pitch accent types are realised through distinct modulations of continuous phonetic parameters.

Pitch excursion is the main phonetic cue used in German to mark information status. Féry & Kügler (2008) found a correspondence between information status and tonal scaling, with narrow focus raising the high tones of pitch accents and givenness lowering them in pre-nuclear position and even cancelling them out in post-nuclear position. Accordingly, the high tones of constituents under narrow focus were consistently higher than those in all-new or given contexts. Similar results are also reported for analyses of several phonetic cues contributing to prosodic marking of focus (e.g., pitch excursion, peak position, duration and accent type in Baumann, Grice, & Steindamm, 2006; Grice, Ritter, Niemann & Roettger, 2017; accent type, duration and articulatory gestures in Hermes, Becker, Mücke, Baumann, & Grice, 2008). Peak position was also found to play a role. Kohler (1991) found that different accent contours were perceived as corresponding to different meanings: late peaks (where F0 rises throughout the stressed syllable, described as L + H*/L*+H in Grice, Baumann & Benzmüller 2005) for emphasis or contrast; medial peaks (where F0 mostly rises throughout the first half of the stressed syllable, described as H*) for new information; and early peaks (where F0 mostly falls throughout the stressed syllable, described as H + L*/H+!H*) for accessible or given information. In particular, a categorical distinction was only found between early and medial peaks, whereas there was a gradient difference between medial and late peaks. In line with these results, Baumann et al. (2006) found a tendency for smaller focus domains (as in contrastive focus) to be produced with later peaks (F0 rising contours).

The above-mentioned studies provide a relatively clear picture of prosodic marking of information status and focus in German, although the mapping is not one-to-one. They demonstrate a probabilistic and gradient relation between information status and accent type (including the absence of an accent).

2.4. A comparative research programme on Italian and German as L1 and L2

One research programme has allowed for the comparison of Italian and West-Germanic languages with an experimental

design that brings to light differences ascribable to language structure: Swerts et al. (2002) and Avesani et al. (2015; 2013) investigated the prosodic marking of information status in Tuscan Italian and Dutch, and in Tuscan Italian and German, respectively. Their findings support Ladd's (1996) observation that Italian strongly disfavours deaccentuation within noun phrases contrary to the pattern in West-Germanic languages.

Swerts et al. (2002) used a card game to semi-spontaneously elicit noun phrases composed of two words (a noun and an adjective), each of which could either be new, given or contrastive according to the context. For Italian, they report an F0 excursion on both words with a hat pattern stretching over the entire noun phrase regardless of the varying information structures, while in Dutch the F0 excursion was on the new element only. It was concluded that Italian fails to deaccent post-focal given elements within noun phrases. A following perception experiment reinforced this finding: Italian listeners could not reliably reconstruct the context in which the noun phrases were produced when listening to them in isolation. This result was replicated in a second perception experiment with the same data (Krahmer & Swerts, 2008), which suggests that these utterances lack sufficient prosodic cues upon which listeners can rely to identify their information status.

Avesani et al. (2015, 2013) successfully replicated the production experiment by Swerts et al. (2002), reporting that in Italian the second word of the noun phrase is always accented, independently of its pragmatic status, whereas the first word can lack an accent in some cases. Interestingly, a range of pitch accents was found for both the first and second words – H*, H + L* and L + H*, but not including the L* found in previous studies on Tuscan Italian (Bocci, 2013). The explanation given is that in Italian, phonological constraints override the mapping between prosody and pragmatic functions, such as focus or information status. In detail, it is argued that the two words of the NP constitute an intonational phrase, whose metrical head at the rightmost position is the stressed syllable of the second word. The first word, being in pre-nuclear position, can optionally be accented, but does not have to be. In contrast, the metrical head (the second word of the NP) has to bear the nuclear accent and the presence of the nuclear accent in the rightmost strong metrical position cannot be overridden by syntactic or pragmatic requirements (Avesani et al., 2015). In German, according to this view, deaccenting of the strongest metrical position of the intonational phrase is allowed and occurs when required by the information status (i.e. in the case of given items).

Avesani et al. (2015, 2013) also extended the investigation to L2 German spoken by Italian learners and L2 Italian spoken by German learners, with the same proficiency level in their respective L2s. They found that Germans successfully reproduced the Italian accentuation pattern, but Italians did not produce the German ones, as only a trace of deaccentuation of given post-focal elements was found in their L2 German (17% of the cases). The authors explain this result with the *Markedness Differential Hypothesis* (Eckman, 1977; as in Rasier & Hiligsmann, 2007 discussed in section 2.1) and the *Similar Differential Rate Hypothesis* (Major & Kim, 1996), according to which marked structures are acquired with a lower speed of learning. This means that Germans can take advantage of possessing both strategies of accentuation, whereas Italian learners of German have a harder task: they have to learn that the distribution of prominences is not necessarily position-based in German and is instead associated with the pragmatic status of the referent. Therefore, the most difficult challenge will be to learn to deaccent the post-focal given referent.

Despite enriching our knowledge of prosodic marking in Italian and German, the studies reviewed above have certain limitations. They investigate a relatively small group of participants, making generalisation of the results difficult, in particular in light of individual differences. They also use a game in which noun phrases are elicited from alternating speakers, where the information status of adjectives and nouns differs across turns. These alternating turns may not have created an engaging interaction between speakers, who may not have assumed the other player's sentences as the context for their own productions, and, instead, speakers may have concentrated on their own list of productions. Finally, the authors focus on the categorical presence or absence of pitch accents and pitch accent type, and only provide limited information on continuous measures – Swerts et al. (2002) report the F0 excursion only. However, previous research on other languages has shown that a closer inspection of continuous phonetic parameters can provide essential information about the expression of pragmatic contrasts (Cangemi, El Zarka, Wehrle, Baumann, & Grice, 2016; Cangemi & Grice, 2016; Cangemi, Krüger, & Grice, 2015; Grice et al., 2017; Mücke & Grice, 2014), raising the question as to whether a closer examination of these parameters might have revealed differences that were not captured in the categorical analysis.

In the following study, we attempt to overcome some of these limitations by a) using a semi-spontaneous interactive board game to elicit different types of information status in a more naturalistic interaction which requires participants to take into account their interlocutor's productions, b) with a larger sample of participants compared to previous studies, and c) performing an innovative acoustic analysis based on periodic energy to investigate how far speakers modulate F0 and prosodic strength to prosodically mark information status.

3. Method

3.1. Participants

40 Italian native speakers, who were learners of German as L2, and 18 German native speakers were recorded. All Italian speakers had grown up in the dialectal area of Naples with parents of the same origin. They were either students at the Goethe Institute (aged between 23 and 65, mean = 33; median = 30; SD = 12.29; 6 females, 4 males), or at the Department of Literary, Linguistic and Comparative Studies (It.: Dipartimento di Studi Letterari, Linguistici e Comparati) at the L'Orientale University in Naples (aged between 19 and 25,

mean = 21; median = 20; SD = 1.2; 27 females, 3 males), with German as a foreign language as one of their main subjects.[1] Italian learners' proficiency levels of L2 German were established based on the language courses they were attending at the time of the recordings and ranged from A2 to C1 CEFR levels (Common European Framework of Reference for Languages – Council of Europe, 2001), in which the notations "A", "B" and "C" correspond to beginner, intermediate and advanced levels of competence, respectively. The proficiency groups resulting from the data collection were unbalanced in number, with only six A- and four C-level learners. Thus, for the sake of a more reliable analysis, learners were recategorised into two proficiency groups, each with a similar number of participants. We defined learners with A2 and B1 levels as beginners and learners with B2 and C1 levels as advanced. This division is not only based on the midpoint of the CEFR proficiency scale, but also on the structure of the reference levels themselves. Indeed, the gap between the abilities required by the B1 and B2 levels is greater than the one between C1 and B2, which makes it a suitable demarcation line for recategorising proficiency levels into two groups only – indeed, the B1 level is also called "Threshold" in the CEFR.

Native German participants came from different dialect areas,[2] but had been living in Cologne for at least three years at the moment of the recording, and were students at the University of Cologne (aged between 22 and 27, mean = 24; median = 24.5; SD = 2.49; 11 females, 7 males). No subject reported to have ever received specific phonetic and/or interactional training, nor to suffer from any speech or hearing problem.

3.2. Data collection

Mono recordings of uncompressed WAV files at 44.1 kHz sample rate and 16-bit depth were collected using headset microphones (AKG C 544 L) connected through an audio interface (Alesis iO2 Express). The two groups of participants described above (section 3.1) were recorded in pairs: Italian learners of German, who were recorded in their native Italian and L2 German, and German native speakers. Participants could self-select their partner for the recordings. Only in a minority of cases in which they could not find a partner, they were matched based on their schedule. In the case of Italian learners of German, the only requirement for self-selecting their partner was the same or a similar proficiency level of L2 German.

Italian participants were recorded at the Goethe Institute in Naples, while German native speakers were recorded at the University of Cologne. During each recording session, two participants sat at two opposite sides of a table without eye contact, which was prevented using a dividing opaque panel. This was done to prevent participants from looking at each other's materials and maximise their use of prosody in the absence of other non-verbal cues.

Prior to the task, participants received written instructions and could ask clarification questions. In the case of learners, the task was first completed in Italian and then repeated in German. Before carrying out the same task in their second language, learners watched video instructions explained by a German native speaker to help them get into the language mode and reduce L1 bias. Finally, all speakers were provided with a sociolinguistic questionnaire.

3.3. Elicitation method

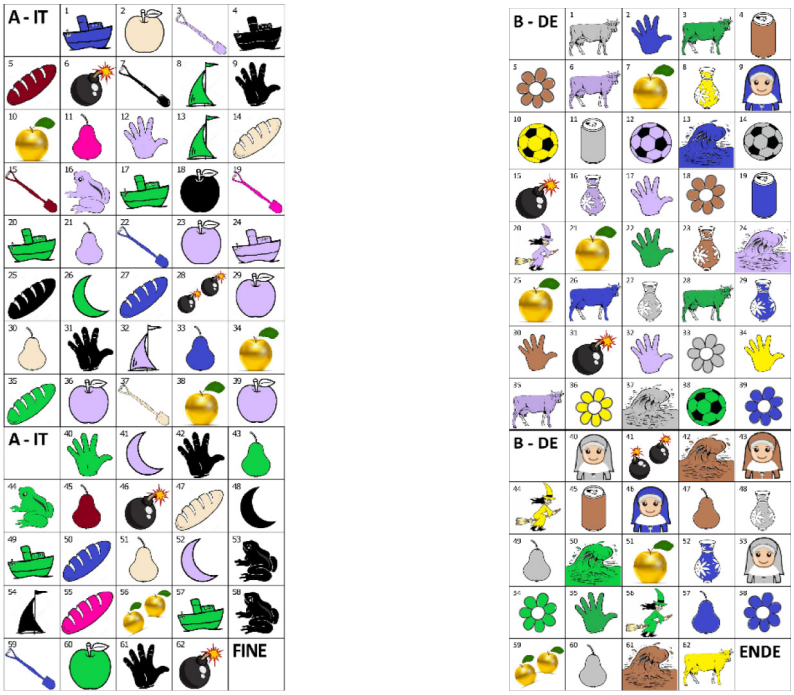
As discussed in section 2.3, some previous studies elicited data using a card game structured in the form of statements between two participants. In that game, both participants receive an equal set of cards containing pictures of different types and varying colours. In alternating turns, one participant picks a card and names its content so that the other participant can align the corresponding card on a board. The two participants alternate the roles of instruction giver and follower. The variation of picture type (noun) and colour (adjective) was designed to create contrastive information statuses in two successive noun phrases produced by participants. However, this type of task presents some disadvantages: it does not favour interaction, so participants might not assume the other's turn as a context for their own statement, and the production of alternating statements might become repetitive and create a list effect. These disadvantages might affect the prosody of participants' realisations and interfere with the pragmatic conditions intended by the experimenters.

Therefore, in the present study special care has been taken in designing an elicitation game to a) increase the degree of interaction and b) avoid the risk of repetitiveness. Despite the difficulty in collecting such specific items in more spontaneous conversation, the design was oriented towards the best compromise possible between ecological validity and the elicitation of noun phrases under the intended pragmatic conditions.

To do so, we created a semi-spontaneous interactive board game to be played in pairs. Each participant received a differently randomised board containing 62 sequentially numbered squares. Each square had a flap which could be lifted to see what is underneath, that is images of various types (nouns) and colours (adjectives). Example boards without the flaps are shown in Fig. 1. All possibly occurring types and colours were listed in the instructions of the game both in visual and written form, but the boards only contained the pictures to avoid interference from reading. Participants were also provided with an additional empty board, displaying only the numbered squares. The task was intended as a distraction, with the aim that participants would pay less attention to their speech. In the instructions, participants were informed that they would go through the table in sequentially alternating turns and that two items are important to win the game, golden apples and bombs. The latter destroys the opponent's golden apples. The person who finds the most golden apples at the end of the game wins provided that they have correctly transcribed the content of the other player's board.

[1] 24 learners had benefited from a stay in a German-speaking country for a variable length of time (from one to ten months), either for a short language course or an exchange period at a partner university. However, the effect of a period abroad is neither accurately quantifiable, nor equal for everyone. It varies according to the amount of input and use of the foreign language (consider, for example, exchange students who do not manage to establish regular contact with locals, or decide not to attend a German language course). Moreover, since a period of immersion in the foreign language contributes to overall language proficiency, we did not consider this variable separately.

[2] Due to the limitations imposed by the COVID-19 pandemic it was not possible to strictly select German native speakers from the same dialect area. In particular, they were born in places above the *Benrath line* (Wenker, 1877) and between the latter and the *Speyer line* (Paul, 2013). Thirteen speakers were from North Rhine-Westphalia, two from Lower Saxony and three from Hesse. However, none of the participants reported a mastery of the dialectal variety of their place of origin.



(a) Italian board for Player A (A-IT) (b) German board for Player B (B-DE)

Fig. 1. Example boards of the elicitation game.

Participants were instructed to sequentially lift the flaps of the table in alternating turns and communicate with a suggested script exemplified in the instructions. The game proceeds as follows: Player A starts uncovering the first picture by lifting the corresponding numbered flap (square number one) and asks player B if they have the picture they see, using a question in which they mention the type and colour of the image (e.g. “Do you have a yellow cow?”). To answer the question, Player B also uncovers the picture in square number one on their own board, and answers with yes or no, followed by mentioning the type and colour of the matching or mismatching image (e.g. “Yes, I have a [matching image]” or “No, I have a [mismatching image]”). This exchange constitutes a game turn. Player B would take the next turn and ask a question about the next square. This alternation of exchanges between players continues until all the squares are revealed. At the end of each turn, both players write on the board (with 62 empty spaces), what their opponent has on their board. At the end of the game, each participant counts how many golden apples they found that were not destroyed by the opponent’s bomb. The person with the most golden apples checks the correctness of their list (the empty board they filled up throughout the game) together with the opponent and wins or loses the game accordingly. This is done in order to keep participants alert and engaged in the game throughout.

Players’ answers about mismatching images by either type or colour, or both type and colour contain our target noun phrases with contrastive elements. The game also elicits yes-replies about matching images by both colour and type. These were inserted only to avoid the bias of a negative answer. Example 1 provides one exchange for each of the possible information status conditions in the Italian and German versions, with the respective English translations. The questions serve as pragmatic context and the replies as carrier sentences of the noun phrases, marked in bold:

(1.a) Italian

A:	*Hai*	*una*	***mano***	***nera?***
	have	a	hand	black
	‘Do you have a **black hand**?’			

B:	*No,*	*ho*	*una*	***mano***	***lilla.***
	no	have	a	hand	lilac
	‘No, I have a **lilac hand**.’				

(1.b) German

A:	*Hast*	*du*	*eine*	***braune***	***Blume?***
	have	you	a	brown	flower
	‘Do you have a **brown flower**?’				
B:	*Nein,*	*ich*	*habe*	*eine*	***braune***	***Nonne.***
	no	I	have	a	brown	nun
	‘No, I have a **brown nun**.’					

The pseudo-randomisation of the sequence of images in the boards followed the criterion that no two identical nouns or adjectives could occur at two subsequent turns to avoid an

unintended degree of activation of the elements deriving from the preceding turn, interfering with the desired prosodic realisations within the turn. With this method, we elicited data from Italian learners of German in both their L1 and L2, and German native speakers.

3.4. Corpus

The target noun phrases were derived from turns in the elicitation game. All NPs are composed of a disyllabic noun and a disyllabic adjective, both with penultimate stress, and correspond to three different information status conditions: given-new (GN), new-new (NN) and new-given (NG).[3] However, we will now report on GN and NG only, since NN and GN presented highly similar realisation in both learners' L1 (see Sbranna et al., 2023 for the complete data analysis in Italian L1) and L2 (Sbranna, 2023). Notice that in Italian the noun precedes the adjective, while in German the adjective precedes the noun. All target NPs analysed here are listed by information status condition and language in Table 1. Each speaker uttered each noun phrase only once.

We do not include in our analysis the functional elements which were merely intended to give the game a goal and avoid repetitiveness, i.e. the mentions of golden apples, bombs and noun phrases inserted in yes-reply carrier sentences, whereby both the noun and the adjective are contextually given.[4] The first two exchanges were used as a training phase and, therefore, also not included in the analysis.

It is important to mention that in some cases, speakers were so engaged in the game that they forgot the suggested script for the interaction. For this reason, some tokens were not realised as prescribed, resulting in a few missing items. However, these cases demonstrate that the game succeeded in engaging the speakers, resulting in more spontaneous behaviour than generally expected from a scripted task. Table 2 contains the number of items for each condition and language group. To reduce potential confirmation biases, we decided not to exclude any items from the acoustic analysis based on the subjective impression of what a "good" or "bad" item is.

3.5. Metrics

For the continuous analysis of prosodic parameters, we employ the open-source *ProPer* workflow (Albert et al. (2020)) to obtain measurements based on periodic energy to account for the shape of F0 and prosodic strength in a quantifiable manner (see overview in Albert 2023:141–158).

ProPer derives the periodic energy curve from *Praat*'s signal processing objects (Boersma & Weenink, 2021) and calculates various metrics to account for the pitch contour trajectories and the prosodic strength of syllabic intervals using *R* (R Core Team, 2021). The current *ProPer* output contains values of three main metrics that are based primarily on the periodic energy curve and its interactions with F0: a) *periodic energy mass* to account for prosodic strength (or *weight*) of syllables, b) *Synchrony* to account for F0 slope within syllables, and c) $\Delta F0$, to account for the difference in F0 between syllables.

Periodic energy is a selective measure of the intensity of the pitch-inducing periodic components of the acoustic signal, reflecting the strength of F0 and its related harmonic partials, which we also link to the linguistic notion of *sonority* (see Albert & Nicenboim, 2022). The periodic energy curve of speech signals is very similar to the amplitude envelope of the general acoustic intensity, and even more so to the amplitude envelope of filtered speech signals that were designed to track syllables (e.g. Port et al., 1996; Galves et al., 2002; Räsänen et al., 2018; Tilsen & Arvaniti 2013). The periodic energy curve exhibits a sequence of fluctuations, with boundaries that are denoted by two subsequent local minima, i.e. onset and offset of syllabic intervals.

The prosodic strength of each periodic energy fluctuation is reflected by the area under the periodic energy curve within syllabic intervals. This metric is the *periodic energy mass*, henceforth *Mass*. The quantitative aspects of this measurement are taken as the integral of duration and power via summing, given that we cannot experience duration or power in isolation (see Roessig et al. 2022 for the success of the Mass metric in predicting the presence of a pitch accent, outperforming both RMS amplitude and vowel duration measurements in that comparison). Mass values are computed with respect to other syllables in the same utterance as a local relationship in the speech signal (a syllable can be 'strong' only in relation to other syllables in its temporal vicinity). Mass is normalized to yield values above '1′ to indicate strong Mass, above the local average, and values below '1′ to indicate weak Mass, below the local average (Fig. 2a).

For each syllabic interval, the Centre of Mass (CoM) is calculated as the average time point within the interval, weighted

Table 1
Target noun phrases by condition in Italian and German.

Condition	Italian	*Gloss*	German	*Gloss*
GN	mano lilla	*lilac hand*	braune Welle	*brown wave*
GN	nave nera	*black ship*	blaue Blume	*blue flower*
GN	mela verde	*green apple*	graue Vase	*grey vase*
GN	rana lilla	*lilac frog*	braune Nonne	*brown nun*
GN	vela nera	*black sail*	blaue Birne	*blue pear*
GN	luna verde	*green moon*	graue Dose	*grey can*
NG	mela lilla	*lilac hand*	graue Welle	*grey wave*
NG	mano nera	*black hand*	braune Blume	*brown flower*
NG	nave verde	*green ship*	blaue Vase	*blue vase*
NG	luna lilla	*lilac moon*	graue Nonne	*grey nun*
NG	rana nera	*black frog*	braune Birne	*brown pear*
NG	vela verde	*green sail*	blaue Dose	*blue can*

Table 2
Amount of noun phrases collected by group and condition.

	Italian L1	German L2	German L1
GN items	222	240	108
NG items	228	239	107
Tot. items	450	479	215

[3] This study is concerned with the dimension of cognitive states (new and given, with all elements being equally accessible since they are listed in the instructions) although in our items this dimension overlaps with pragmatic functions given the context of elicitation, i.e. NN with broad focus/two contrastive foci, and the new element of GN and NG with contrastive focus.

[4] The reason for not including this latter given-given condition (GG) is that due to the semi-spontaneous nature of the game, many speakers simply answered "yes" to the context question without following the suggested script and using the noun phrase, in other words, we consider these productions driven by the script only.

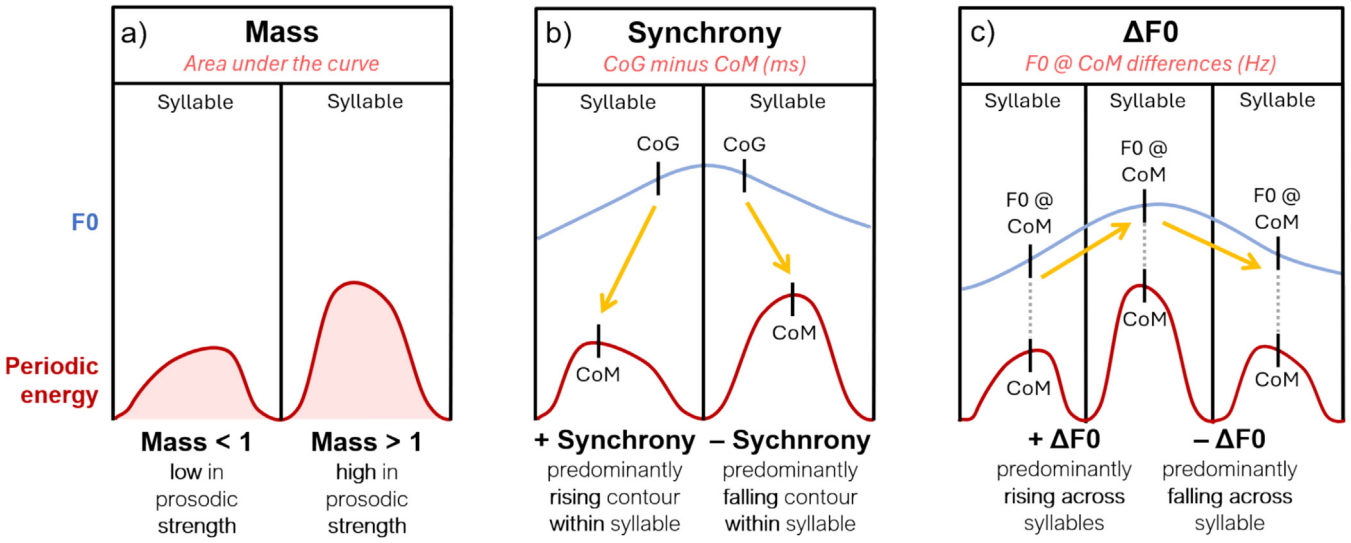


Fig. 2. ProPer metrics: Schematic descriptions and interpretations.

by periodic energy (see Cangemi et al., 2019 for details). Likewise, the Centre of Gravity (CoG) is calculated as the average time point within the same interval, weighted by F0 (derived from the *Tonal Centre of Gravity* in Barnes, Veilleux, Brugos, & Shattuck-Hufnagel, 2012). The CoG and CoM are sensitive to the continuous and non-linear aspects of the contour shape, and the distance between them within intervals is indicative of the overall F0 shape in the given syllable, yielding the metric termed *Synchrony*: more negative values reflect steeper falling contours, more positive values reflect steeper rising contours, and values closer to zero reflect symmetric contours, very often a plateau (Fig. 2b).

ΔF0 is a measurement of the change in the F0 trajectory across intervals. It is achieved by calculating the difference in F0 value at the CoM in a given interval relative to the previously occurring interval. This yields an indication of F0 movement and excursion across the two adjacent intervals, i.e. higher for positive values, lower for negative values, or unchanged for values around zero (Fig. 2c). Notice that in our analysis, the value of ΔF0 on the first syllable does not refer to the difference from the previous portion of the utterance, which we do not take into account, but rather the difference from the median F0 of that speaker. This value is mostly useful to flag cases starting with a relatively high F0, whereby the first syllable is expected to be higher than the median, resulting in a positive ΔF0 value.[5]

In sum, Synchrony reflects the shape of the F0 contour within syllabic units (akin to *tonal peak alignment*; e.g. Arvaniti, Ladd, & Mennen, 2006), while ΔF0 reflects the shape of the F0 contour across syllabic units (more comparable to common *scaling* measurements). These two characterisations of the *quality* of the pitch contour in terms of pitch height are complementary to the Mass metric that characterises the *quantity* of the sonorous material and perceived pitch in one variable (conflating the abstract duration and power dimensions).

We normalise Synchrony values by computing raw Synchrony (in absolute milliseconds) relative to the length of the containing interval (yielding a percentage-based scale). Likewise, we normalise ΔF0 values by computing the raw ΔF0 (in Hz) relative to the speaker's F0 range (again, yielding a percentage-based scale). Finally, Mass values are normalised by computing all the raw Mass values within a given utterance relative to each other (yielding a scale in which '1′ stands for the local average and '0′ stands for the minimum). We apply this workflow to the acoustic analysis of the target noun phrases. In particular, we use the periodic energy curve to produce *Periograms*, ProPer's enriched visual representations of F0 trajectories modulated by periodic energy (Albert et al., 2018), and process the derived ProPer measures further in R to display distributions of our metrics and perform statistical inference analyses. Examples of single utterances from the experimental dataset, with ProPer landmarks and metrics overlaid, are available in Fig. 3 below.

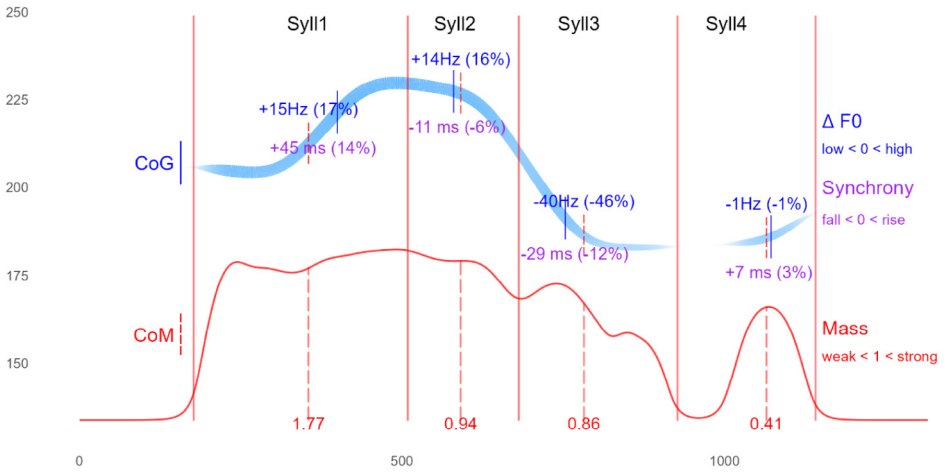
3.6. Bayesian inference

We fitted Bayesian hierarchical linear models using the Stan modelling language (Carpenter et al., 2017) and the package *brms* (Bürkner, 2016) to examine the robustness of the differences found across languages and conditions.

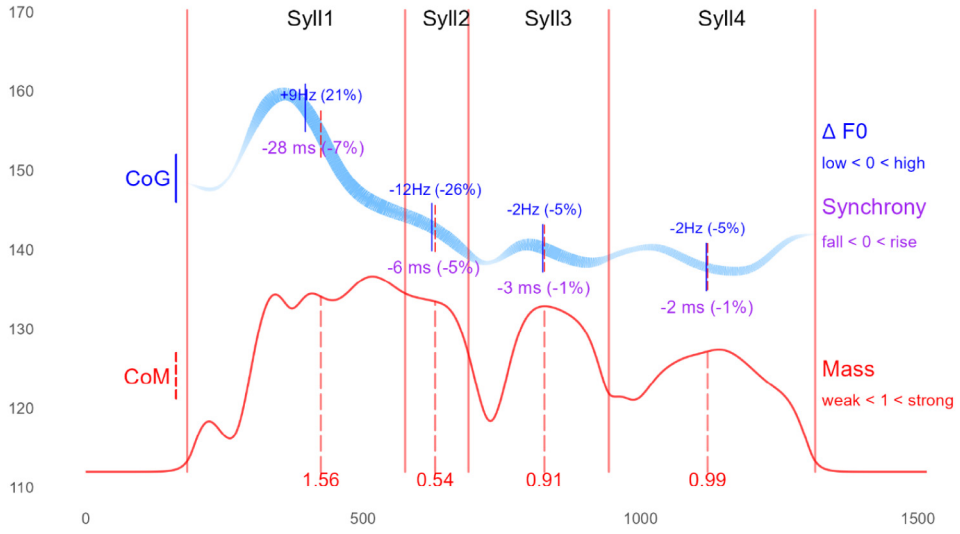
For each language group, the differences among conditions in Synchrony, ΔF0 and Mass were tested as a function of factors CONDITION (reference level "NG"), SYLLABLE (reference level "Syllable 1") and their interaction. As random effects, the models include random intercepts for TOKEN and SPEAKER. For SPEAKER the models also include by-speaker random slopes for CONDITION and SYLLABLE and correlation terms between all random effect components. For models testing the differences across groups, the fixed effect GROUP[6] was added to CONDITION and SYLLABLE, as well

[5] We use this method on the first syllable of the target NPs instead of ΔF0 referring to the previous portion of the utterance because our data often display a pause between the carrier sentence ("No, I have a [. . .]") and the target item, i.e. the noun phrase. Furthermore, since we do not always have analysable material preceding the target item this choice allows us to present a unified measurement for ΔF0 values in Syllable 1 across all data.

[6] The factor group had three levels including ITALIAN L1, GERMAN L1 and LEARNERS as a unique group. This was done after having checked for statistical differences between beginners and advanced learners as described in the result section.



(a) GN. The token is *graue Vase*



(b) NG. The token is *braune Dose*

Fig. 3. Example periograms displaying F0 (blue curve at the top half) and periodic energy (red curve at the bottom half) for two tokens in two conditions, produced by Italian learners of German as L2. Note how the thickness and darkness of the blue F0 curves are modulated by the corresponding red periodic energy curves below them. Solid red vertical lines denote interval boundaries, dashed red vertical lines denote the centre of periodic energy mass (CoM) and short blue vertical lines on the F0 trajectory denote the centre of gravity (CoG) of the F0 curve. Numeric values per syllable are colour-coded with the three ProPer metric variables, ΔF0 (blue), Synchrony (lilac) and Mass (red). Values on the x-axis show time in milliseconds, and values on the y-axis show frequency in Herz (relevant only to the F0 curve). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

as a three-way interaction between them.nThe models include the whole corpus, even though we report here only on the contrasts relevant to NG and GN conditions.

Details about prior choice, model settings, results and posteriors can be found in the *RMarkdown* file at the Open Science Framework (OSF) repository.[7] There was no indication of convergence issues (no divergent transitions after warm-up; all Rhat = 1.0), including from visual inspection of the posterior distributions.[8]

For all relevant contrasts (δ), we report the expected values under the posterior distribution and their 95% credible intervals (CIs), that is the range within which an effect is expected to fall with a probability of 95%. To ensure comparability with conventional null-hypothesis significance testing, we also report the posterior probability that a difference between contrasts is bigger than zero ($\delta > 0$). In particular, we assume that there is (robust) evidence for a hypothesis that states $\delta > 0$ if zero is (by a reasonably clear margin) not included in the 95% CI of δ and the posterior $P(\delta > 0)$ is close to one (cf. Franke & Roettger, 2019).

4. Results

We will first describe results for L2 learners' realisations as a whole group and across proficiency levels, and

[7] See "Part II – Prosodic Marking of Information Status" at https://osf.io/9ca6m/?view_only=d3406b6d6bdc43e39dabbbca7319eff9.

[8] The model's assumption for ΔF0 was not fully satisfied since our posterior simulations are less leptokurtotic than our actual data. Still, the model does not show convergence problems.

then compare them to their native and target language. For conciseness, we will refer to the syllables of the noun phrases with the labels used in the figures, Syll1–Syll4: Syll1 and Syll2 comprise the first word, Syll3 and Syll4 comprise the second word; the stressed syllables are Syll1 and Syll3 (all words are disyllabic with a penultimate/trochaic stress pattern).

4.1. Learners' interlanguage: L2 German

Fig. 3 displays two example periograms for the two information status conditions, showing two different F0 patterns across conditions. In the GN condition, F0 is rising throughout Syll1, reaches a peak on Syll2 and finally falls on Syll3. In NG the peak is fully realised on Syll1, where the falling movement already begins taking place. The presence of new vs. given information in the final position of the noun phrase is characterised by two different F0 patterns distinguished by the timing of the falling F0 movement: it is fully realised on the stressed syllable of the second word (Syll3) in GN, or on the stressed syllable of the first word (Syll1) in NG.

The aggregated data of Synchrony, ΔF0 and Mass in Fig. 4 confirm these observations. Specifically:

- *Synchrony*. Mostly positive distribution on Syll1 of GN reflects a rising F0 trajectory. In contrast, more negative values on Syll1 reflect the earlier F0 fall in NG. The effect of the later F0 fall in GN is reflected in more negative Synchrony values on Syll3 than in NG.
- *ΔF0*. Mostly positive values on Syll2 and negative values on Syll3 are indicative of the F0 peak on Syll2 in GN. In contrast, mostly negative values already on Syll2 in NG indicate that the F0 peak is located within Syll1, resulting in less negative values on Syll3 as compared to GN.
- *Mass*. All conditions present the same pattern and display strong energy on Syll1 and weaker energy on Syll3 (although slightly stronger than Syll4). This pattern does not reflect information status contrasts between NG and GN.

To summarise, we found two different F0 contours distinguished by the location of the F0 falling movement: the bulk of this falling contour is within the stressed syllable of the first word in NG (Syll1) and later within the stressed syllable of the second word (Syll3) in GN, resulting in different values of Synchrony (at Syll1 and Syll3) and ΔF0 (at Syll2 and Syll3). We found one pattern of modulation for Mass across all conditions, namely strong Mass on the first word and weak Mass on the second word (Syll1 vs. Syll3). Thus, we will not test any contrast relative to this metric. We will test Synchrony and ΔF0 contrasts across conditions as follows:

a) *Synchrony*. GN higher than NG on Syll1; GN lower than NG on Syll3.
b) *ΔF0*. GN higher than NG on Syll2; GN lower than NG on Syll3.

Results of the Bayesian analysis support our observations:

a) *Synchrony* on Syll1 is higher in GN (δ = 4.44, CI [3.39; 5.62], P (δ > 0) = 1) than in NG. In contrast, Synchrony on Syll3 is lower in GN (δ = 1.62, CI [0.52; 2.78], P (δ > 0) = 0.99) compared to NG.
b) *ΔF0* on Syll2 is higher in GN (δ = 19.52, CI [13.83; 25.73], P (δ > 0) = 1) than in NG. In contrast, ΔF0 on Syll3 is lower in GN (δ = 7.15, CI [2.99; 11.66], P (δ > 0) = 0.99) than in NG.

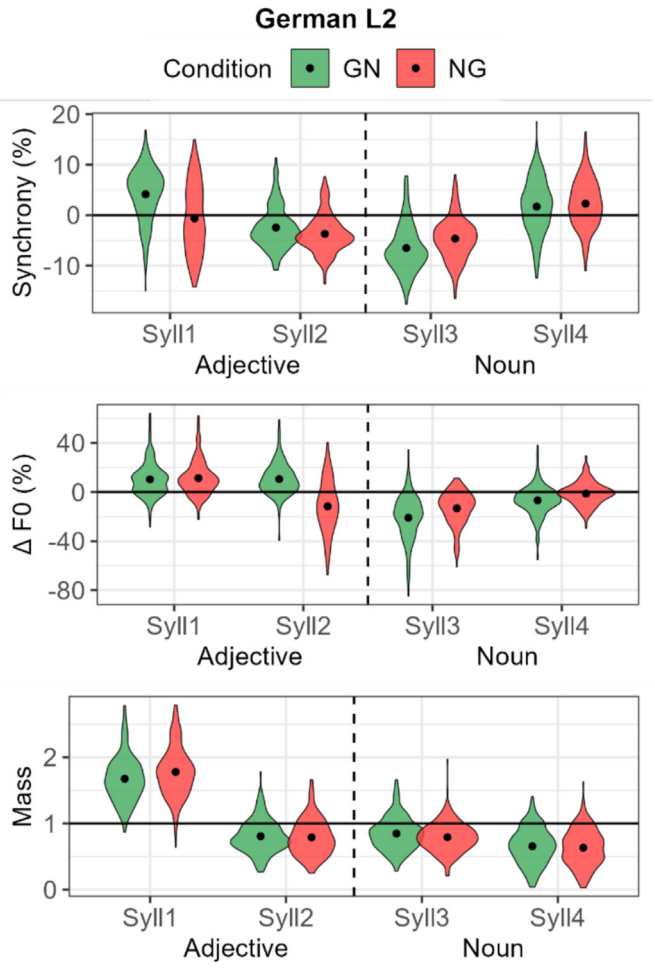


Fig. 4. Aggregated values of Synchrony, ΔF0 and Mass pooled across German L2 learners. The x-axis displays the four syllables of the noun phrases, with syll1 and syll2 being the adjective and syll3 and syll4 the noun. Information status conditions are colour-coded: green for given-new (GN) and red for new-given (NG). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

To summarise, learners distinguish two intonation patterns through the modulation of the F0 falling movement on the first word (similarly to what has been found in their native language in Sbranna et al., 2023). In detail, they use a falling F0 contour on Syll1 to mark the NG condition, whilst they realise the F0 falling contour mostly within Syll3 in the GN condition. Contrary to F0, the modulation of Mass across conditions does not result in distinct strength patterns. Learners use Mass in a way that is unexpected with respect to the literature about both L1 Italian and L1 German, with stronger Mass on Syll1 compared to Syll3, independently of information status.

4.1.1. Analysis by proficiency level

As mentioned in section 3.2, learners were categorised into two proficiency levels, i.e. beginner and advanced learners. However, both proficiency groups produce the same patterns for F0 and Mass modulation. Only subtle differences across proficiency levels can be spotted, with beginners displaying less discrimination between conditions as compared to advanced learners. Since this distinction was not found to be of great importance, the following description only highlights

the results. The full account on proficiency level is available in the online supplementary materials.[9]

The models suggest that most of the cross-group differences concern the modulation of Synchrony and ΔF0 values on the first three syllables. These are now displayed in the same two-dimensional space in Fig. 5, where advanced learners seem to show less overlap and more distinct clusters separating the two conditions.

The analysis by proficiency level has shown that robust differences between beginner and advanced learners are present, but that they are mostly very small in magnitude. Specifically, the advanced group seems to enhance the difference between their NG and GN realisations more when compared to beginners, by modulating synchrony and ΔF0 towards opposite directions. In contrast, beginners consistently differentiate the two conditions less clearly on the second word by neutralising F0 and mass distinctions across conditions. Still, these differences do not result in overall distinct patterns, and the two proficiency levels show the same general trends for the two pragmatic conditions. For this reason, we will consider them as one single group in the next section, i.e. when comparing L2 German to L1 Italian and L1 German.

4.2. Interlanguage compared to learners' native and target languages

We performed the same ProPer analysis for the Italian L1 (see Sbranna et al., 2023) and German L1 data (Sbranna, 2023) and tested contrasts across conditions and groups statistically. To describe data more easily and concisely, we show here the averaged F0 contours and the values for Synchrony, ΔF0 and Mass for syllables in the comparison.

4.2.1. F0 contours

Fig. 6 displays the averaged F0 contours found for the realisation of NG and GN in the three language groups, i.e. Italian L1, German L2 and German L1.

Comparing first learners' native (L1 Italian) and target language (L1 German), it is apparent that the two languages use different strategies to express the contrast between the two pragmatic conditions: L1 Italian prosodically differentiates the two conditions mostly on the first word, while L1 German does so mostly on the second word of the noun phrase. Specifically, in L1 Italian, NG shows an F0 fall early in the first syllable while in GN the falling trajectory begins on the second syllable. In L1 German, in contrast, NG is more similar to Italian GN, where the F0 fall begins on the second syllable. The GN contour exhibits a "hat pattern", with high F0 stretching into the second word such that the bulk of distinctive F0 fall begins relatively late on the third syllable. Thus, from an Italian learner's perspective, the timing of the F0 fall in German GN is later than in their native baseline.

Based on these comparisons, learners seem to transfer their L1 F0 shapes to their L2 realisations, as the intonation patterns in the L1 and L2 appear to be quite similar,[10] even though they exploit a reduced F0 range as compared to their baseline (more similarly to the target language, which will be further addressed in the final discussion). In other words, as in Italian L1, the German L2 contours do distinguish information status within noun phrases, but they do so using a different timing of the F0 fall on the first word and not differentiating F0 on the second word, as in L1 German. Thus, although they differentiate between the two conditions, their productions do not match the target contours. In NG, the learners' F0 fall takes place too early in the first word compared to the target (on the first syllable instead of the second syllable) and, in GN, learners do not produce a hat pattern across the two words and the F0 fall occurs before the new element (on the second syllable of the first word, instead of the first syllable of the second word). As a result, differences between learners' realisations and their target language are evident on the first word in the NG condition and on the second word in the GN condition.

Fig. 7 provides values derived from the F0 modulation across the three language groups, including Synchrony and ΔF0 for the relevant locations in the noun phrases in light of the differences observed, i.e. Syll1 and Syll3 for Synchrony, and Syll2 and Syll3 for ΔF0. Distributions and mean values of Synchrony and ΔF0 for the learner group are often midway between their native and target languages. This is indicative of the above-mentioned reduction in F0 range because values of Synchrony and ΔF0 approaching zero indicate less pronounced changes within and across syllables. Therefore, it appears that learners do not completely transfer the pattern from their L1 but, instead, do change something in their L2 productions as compared to their native language, despite not achieving entirely different F0 contours. These observations will be tested as follows:

a) *Synchrony*. In L2: NG lower than both native groups on Syll1; GN higher than L1 Italian and lower than L1 German on Syll3.
b) *ΔF0*. In L2: NG higher than L1 Italian and lower than L1 German on Syll2; GN higher than L1 Italian and lower than L1 German on Syll3.

Results of the Bayesian analysis support our observations:

a) *Synchrony*. On Syll1 in NG, values for the L2 group are lower than both in the native (δ = 1.09, CI [0.31; 1.81], P(δ > 0) = 0.99) and the target language (δ = 1.30, CI: [0.29; 2.19], P(δ > 0) = 0.99). On Syll3 in GN, values for the L2 group are higher than in the native (δ = 2.66, CI [1.83; 3.43], P(δ > 0) = 1) and lower than in the target language (δ = 6.02, CI: [5.02; 7.04], P(δ > 0) = 1).
b) *ΔF0*. On Syll2 in NG, values for the L2 group are higher than in the native language (δ = 5.01, CI [2.3; 7.77], P (δ > 0) = 1), but still lower than in the target language (δ = 3.49, CI [0.03; 7.01], P (δ > 0) = 0.97). The same holds for Syll3 in GN (difference to L1 Italian: δ = 13.89, CI [11.57; 16.24], P (δ > 0) = 1; difference to L1 German: δ = 16.56, CI [13.46; 19.81], P (δ > 0) = 1).

4.2.2. Mass

Mass reflects prosodic strength as it is derived from a calculation accounting for power and duration, two parameters which are often involved in prosodic highlighting/attenuation together with F0 movement and can result in the perception of the presence or absence of an accent. We will, therefore, discuss values of Mass across language groups for the third syllable, as this is the one which should differ across the two

[9] See section 7.1 "Analysis by proficiency level" of the.pdf file at https://osf.io/9ca6m/?view_only=d3406b6d6bdc43e39dabbbca7319eff9.

[10] Notice that the difference in duration of the first syllable across languages depends on the underlying segmental material, i.e. a different syllabic structure. At the beginning of German NPs there are adjectives like "graue" and "blaue" with two consonants in the onset and a diphthong as the nucleus, while in Italian all syllables are composed of a single consonant and a vowel.

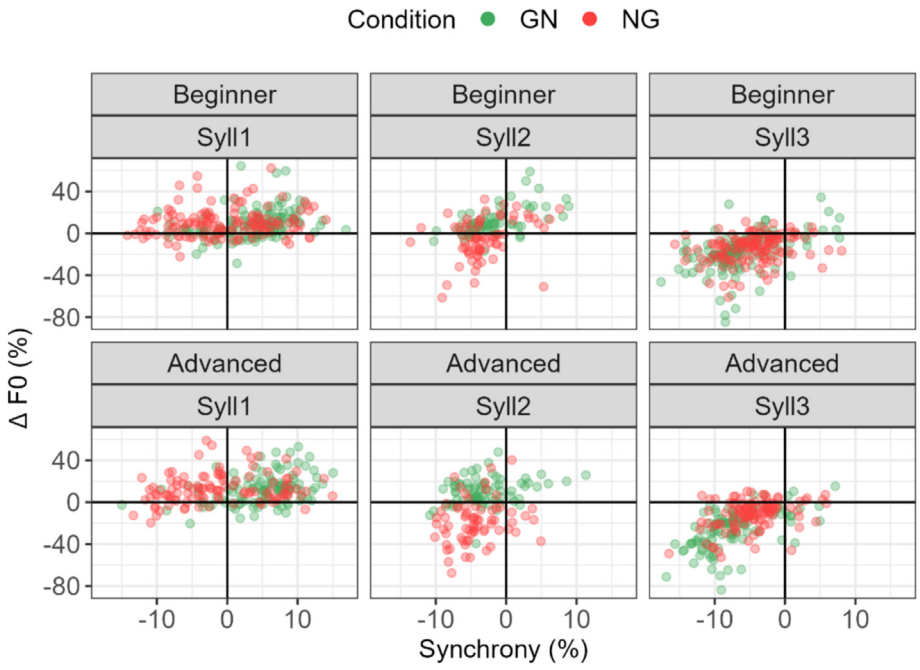


Fig. 5. Synchrony and ΔF0 values for syllables one to three across proficiency levels. Values of Synchrony are displayed on the x-axes, values of ΔF0 on the y-axes. Information status conditions are colour-coded: green for given-new (GN) and red for new-given (NG). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

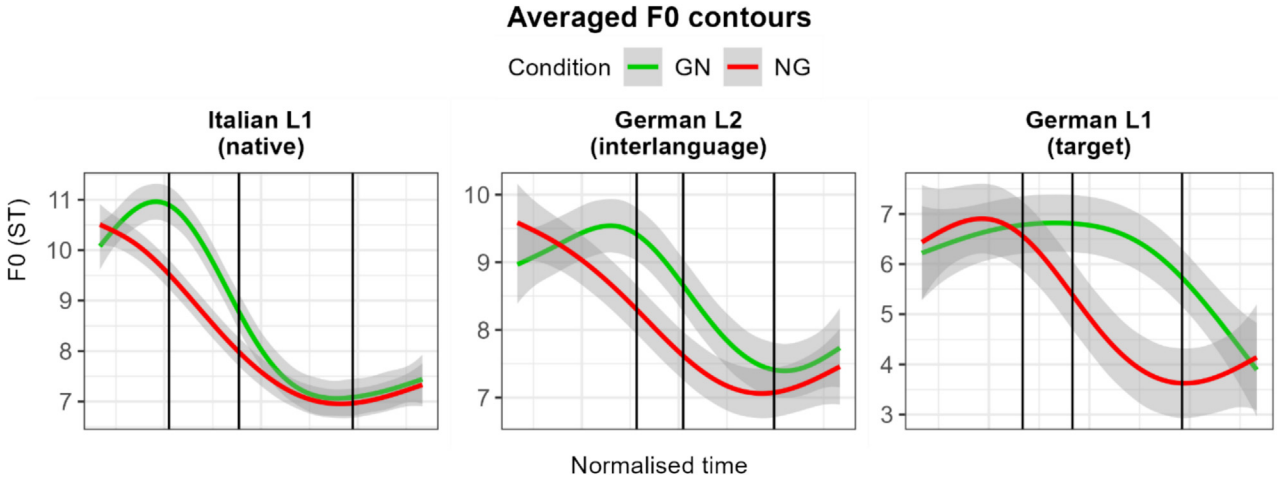


Fig. 6. Averaged F0 contours pooled across speakers for each language group. The y-axis shows F0 in semitones, while the x-axis shows normalised time (across items for each of the four syllables separately) aligned at the boundary between the two words of the noun phrase. Syllable boundaries of the noun phrase are marked by vertical black lines. The grey area around the contours represents the standard error and contours are colour-coded according to information status: green for given-new (GN) and red for new-given (NG). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

L1s according to the literature, with Italian expected to always accent it in accordance with phonological positional rules, and German expected to only accent post-focal information when it is new, and instead to deaccent it when it is given.

Values of Mass for the third syllable of the noun phrase are shown in Fig. 8. Looking first at native Italian and native German, we see that our results provide evidence in line with the literature (section 2.1 for Italian and section 2.2 for German). L1 Italian displays strong Mass on Syll3, both when it is new and when it is given and post-focal, supporting the finding that the final word requires an accent in the language, independently of pragmatic status. By contrast, in L1 German we only observe strong Mass on Syll3 when it is new, whereas Mass is weak on the post-focal given element, showing that, in line with previous studies reporting deaccentuation, the final position is prosodically highlighted or attenuated according to information status.

As previously shown, in their L2, Italian learners clearly present weak Mass across conditions on Syll3, with values similarly distributed below one. The values appear to be even more negatively distributed than in the NG condition by L1 German speakers, also displaying weak Mass. As a result, learners do not seem to transfer Mass patterns from their native language as they do for F0 and, instead, show prosodic attenuation as in the target language, which might be interpreted as an attempt to reproduce deaccentuation. However, in compar-

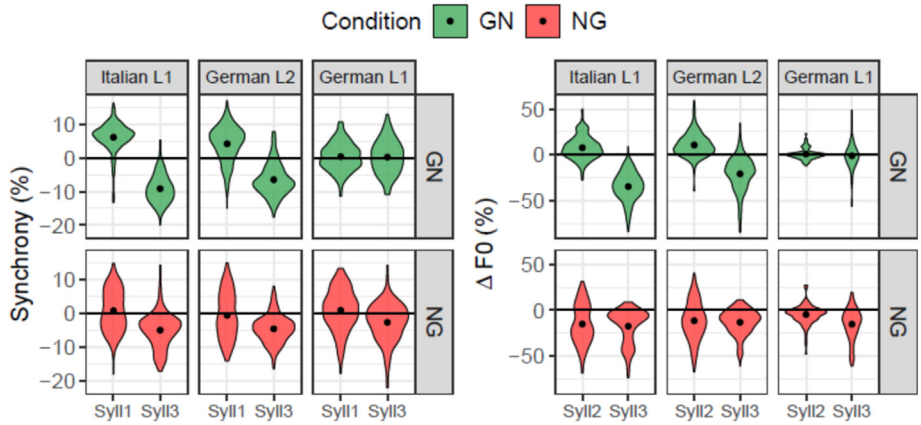


Fig. 7. Aggregated values for Synchrony and ΔF0 for the relevant syllables across language groups. Synchrony values are displayed for Syll1 and Syll3 (left panels), while ΔF0 values are shown for Syll2 and Syll3 (right panels). Information status conditions are colour-coded and positioned on two separate rows: green for given-new (GN, upper row) and red for new-given (NG, bottom row). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

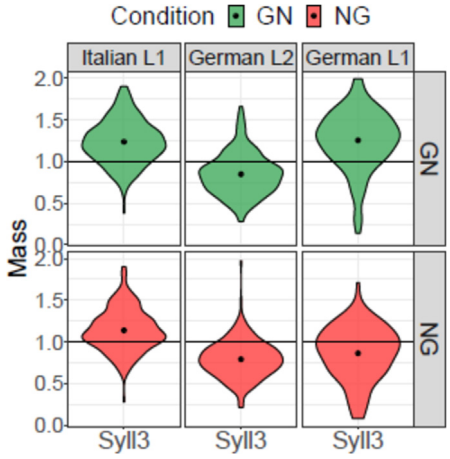


Fig. 8. Aggregated values of Mass for syllable three across language groups. Information status conditions are colour-coded and positioned on two separate rows: green for given-new (GN, upper row with syllable three being a new item) and red for new-given (NG, bottom row with syllable three being a given item). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

ison with L1 German, learners' Mass on Syll3 is weak not only in NG but also in GN, i.e. they appear to deaccent new information as well.

To summarise, learners' prosodic strength patterns diverge from their native Italian and tend to reproduce the native German for the NG condition, with the difference that a) they seem to attenuate the post-focal element even more than native speakers of German, and b) this same pattern is extended to GN as well, showing that they do not use prosodic strength to mark information status contrasts (which is something they do not do in their native Italian either). This will be tested as:

a) *Mass*. In L2: NG and GN on Syll3 lower than in either native language.

The results of the Bayesian analysis support our observations.

a) *Mass*. On Syll3 in both conditions, the L2 group presents lower values than both in the native (difference for NG: δ = 0.34, CI [0.29; 0.39], P (δ > 0) = 1; and GN: δ = 0.38, CI [0.33; 0.43], P (δ > 0) = 1) and target language (difference for NG: δ = 0.09, CI [0.03; 0.15], P (δ > 0) = 0.99; and GN: δ = 0.42, CI [0.35; 0.48], P (δ > 0) = 1).

4.2.3. Summary

To summarise, Italian learners of German produce F0 shapes that resemble those in their L1, but they use a different Mass pattern across all conditions, one that is similar to L1 German deaccentuation in the NG condition. However, modelling the phonetic details reveals that learners' modulations of both F0 and prosodic strength are robustly different from both their native Italian baseline and the native German target. In particular:

- Comparing F0 shapes in L2 German and L1 Italian, the models provide evidence for learners using more negative Synchrony on Syll1 in NG when speaking their L2 German than their L1 Italian, which suggests that they used a less steep slope in the early F0 fall in L2. Likewise, the ΔF0 values in L2 were closer to zero across conditions. This suggests that learners tend to exploit a narrower F0 range in their interlanguage. These consistent differences are, however, slight, as can be informally observed in Fig. 8, and are unlikely to have a communicative value.
- Comparing the patterns of Mass in L2 German and L1 German in the NG condition, the model shows compelling evidence for more negative Mass on the first syllable of the given item when this item is second (Syll3 in NG) in the interlanguage, meaning that learners prosodically attenuate this given word even more than native German speakers. This might suggest that learners perceive a reduction in Mass as a salient cue to be reproduced, intending to sound more like a native German when speaking their L2 and possibly overgeneralising its use.

5. Discussion

Using an exploratory bottom-up approach, we found that Italian learners of German transfer aspects of their L1 intonation contours. However, statistical results indicating systematically different modulations of F0 for the two conditions provide evidence that the transfer is only partial. Specifically, learners produce the fall in the NG condition with a shallower slope and a reduced F0 range overall. Although a narrower F0 range is characteristic of L1 German, it is not very likely that learners intentionally compress their Italian native F0 range to approach

the target language. In fact, other L2 studies involving different language pairs have made the same observation about reduced F0 ranges (Dutch learners of Greek in Mennen, 1998; Taiwan Mandarin learners of English in Visceglia, Tseng, Su, & Huang, 2011; Italian learners of English in Urbani, 2012; French learners of German and German learners of French in Zimmerer, Jügler, Andreeva, Möbius, & Trouvain, 2014; Chinese learners of Japanese in Shi, Zhang, & Xie, 2014). This has been interpreted as a characteristic of interlanguages, not necessarily ascribable to transfer or learning effects. Instead it is more likely to be related to insecurity in speaking a second language (Mennen, 1998; Shi et al., 2014; Zimmerer et al., 2014). This is further corroborated by the fine-grained variation we found across proficiency levels, where beginners compressed their F0 range more than advanced learners.

We also found that, as in their L1 Italian, these learners do not make use of prosodic strength to distinguish the two pragmatic conditions. However, the strength pattern learners use is not present in their L1 as they strengthen the first word and weaken the second, similarly to L1 German productions in the NG condition. This might be an attempt to reproduce a pattern typical of German. However, learners attenuate the second word even more than L1 German speakers, probably as a form of hypercorrection (cf. Eckman, Iverson, & Song, 2013; Kelly, 2022; Petrov, 2021), and apply this pattern across all pragmatic conditions (i.e. also in GN), risking a misinterpretation of its function. This misinterpretation is possibly because in their L1 Italian prosodic strength is not used to encode information status. In Italian, positional rules prevail. This involves accenting words in final position independently of information status. Thus, Italian learners of German may not be aware of the different function it has in German. In an earlier study on the prosody of Italian learners of German, Avesani et al. (2015) also reported cases of prosodic attenuation that were not linked to discourse-related properties, although these accounted for less than a fifth of their data. Our results in line with Avesani et al.'s results, although, for the specific dimension of prosodic strength (as opposed to pitch contours) we found attenuation to be widespread. Both Avesani's and our results appear to be in line with the Markedness Differential Hypothesis (see section 2.1), in that they point to partial acquisition of superficial form without linking this form to information structure in the target language.

For Italian learners of German, we also found a minority of NG items realised in the same way as GN, that is, without matching prosody to the information status of the last element. This phenomenon was also observed in their native language (Sbranna et al, 2023) and assessed through a perception experiment in which Italian native listeners were asked to indicate which information structure corresponded to the contours. Results showed that those NG instances produced with a contour similar to GN (for which the fall on the second syllable of the first word is decisive), were not recognised as such and were confused with GN. A future, in-depth investigation of this minority of NG instances might be useful to observe and verify whether their distribution reveals a link to L2 proficiency.

This subset of NG instances can be explained by at least two factors. The more general explanation resides in the distributional property of the prosodic mapping onto information status, by which no one-to-one mapping exists, but different options are available and intonational strategies can differ in their frequency of occurrence. A complementary context-specific explanation is related to some limitations of this elicitation method. In the context of the elicited noun phrases, prosody is not always necessary for the correct interpretation of the sentence, given the inherent redundancy in the communicative situation. Secondly, speakers may not have paid attention to the question posed by the interlocutor, since game turns are repetitive in their structure, an effect that cannot be completely avoided, despite being limited by the insertion of distractors in the design. Finally, speakers may choose different strategies for accomplishing a task. Instead of a listener-oriented strategy (also referred to as *audience design*; Bell, 1984), where speakers try to make sure that the interlocutor can receive and interpret the message properly, some speakers may have applied a self-oriented strategy. A prerogative to win the board game was to correctly write down all the images named by the interlocutor. As a result, some speakers may have shifted their attention from the interlocutor's questions to the writing task and, consequently, produced their answers without having the pragmatic context of the questions in mind.

Furthermore, although the interactivity of the elicitation method was enhanced in relation to the previous studies, the type of speech investigated here cannot be described as completely spontaneous. Future research should strive for the best possible compromise between spontaneity and systematic data collection, aiming to collect fully spontaneous conversational speech in order to increase the ecological validity of research findings. This is also important in terms of learners' skill assessment, as improvement in controlled exercises or tasks might not be retained in spontaneous speech in which there is a different cognitive load.

We also note a limitation regarding the way learners' proficiency is treated. To allow for more reliable statistical testing, learners were categorised into two main proficiency groups, since the sample contained more intermediate than beginner and advanced learners (according to CEFR classification). To investigate the process of second language acquisition and critically discuss it in relation to the CEFR and its descriptors, a larger sample for each proficiency level would be required. In this way, it would be possible to observe if learners fulfil the skill descriptors related to each level and, if not, allowing a teacher to intervene with apposite and efficient pedagogical tools. Moreover, it should be remembered that proficiency is a continuum upon which the CEFR levels are developed, and looking at single performances within a certain level will surely show individual-specific variability (see discussion on the CEFR labels within a dynamic system approach to second language acquisition in Lowie, 2013). Ideally, future studies would collect longitudinal data and shed light on the individual trajectories of L2 development.

The results presented here have managed to shed new light on interlanguage phenomena due to the methodologies we present in this work, both for data elicitation and for data analysis. Speech style and degree of spontaneity have been shown to have a considerable influence on intonation (for the difference between read and spontaneous speech in Italian varieties and on German, see Grice, Savino, & Refice, 1997;

De Ruiter, 2015 respectively). The design of the current elicitation game was intended to create a high level of interaction between participants, as well as engagement in the task which, together with the lack of eye-contact, appeared to succeed in promoting the use of prosody for conveying different pragmatic meanings (further suggested by perceptual results in Sbranna et al., 2023).

For the data analysis, the ProPer metrics allowed us to compare entire structures with aggregated data, revealing patterns which have not emerged previously in analyses based on discretised categories. Labelling phonological categories entails some degree of subjectivity due to the annotator-specific perception of meaning and expectations based on the native language of the annotator (Röhr, Baumann & Grice, 2022; Cangemi & Grice, 2016; Bishop, 2013). Different annotators can make different choices, and the individual-specific bias is even more problematic when labelling an L2 as it is not a stable system, for which there would be clear and stable reference values for categories (i.e. there is no annotation system dedicated to the interlanguage). Traditional categorical descriptions are useful and necessary for generalisation purposes and SLA research applications (as when creating pedagogical tools for non-specialist teachers and learners), but they are not optimal for exploratory studies that attempt to capture a snapshot of the learners' unique language system, which is a moving target, in a constant state of flux.

We should note that some recent annotation systems are capable of reflecting graded levels of prominence. These include the Rapid Prosody Transcription (Cole & Shattuck-Hufnagel, 2016) which takes the aggregated responses of multiple annotators to present prominence scores, and the DIMA system (Kügler et al., 2015), which uses multiple levels of prominence. However, these annotation systems would have been sub-optimal for our research questions for the following reasons: (1) both systems focus on privileged positions (accented syllables/words and edges of prosodic constituents) so they do not provide information about entire structures; (2) the limited inventory of symbols that describe accents in DIMA (essentially H and L) is not designed to capture subtle differences in pitch contours; (3) RPT provides no direct information on perceived pitch contours.

6. Conclusion

This investigation was inspired by previous research finding that Italians learners of West Germanic languages do not mark information status prosodically within noun phrases, neither in their L1 nor their L2, which was also confirmed in perception. Our goal was to find out whether a close inspection of continuous phonetic parameters would bring to light subtle prosodic modulations for differentiating information status. We found that Italian learners of German do differentiate the two pragmatic conditions, using two F0 shapes which highly resemble their native Italian ones. Still, they do not exhibit a complete transfer and show small, but systematically different modulations of F0 across pragmatic conditions, which are somewhat clearer at an advanced level. Furthermore, we found an interesting pattern when we observed the modulation of syllabic strength in terms of Mass. Learners seem to systematically reduce the strength (in the acoustic power and duration dimensions) of the second word in the noun phrase, regardless of condition (i.e. irrespective of pragmatic considerations of information status).

The present study has also demonstrated promising tools for pedagogical applications.[11] By directly modeling the shape of pitch contours on syllabic anchors (using the Synchrony and ΔF0 metrics), and separately accounting for the strength (or weight) of syllables in the structure (using the Mass metric) we can gain insights into the differences between the prosody of a native language and a target language, as well as their interlanguage. Assessing learning outcomes of aspects of prosody with the ProPer toolbox can help directing teachers towards the strengths and weaknesses of L2 learners in a beneficial way (see Möking et al. 2024 for more direct evidence on the pedagogical applications of our methodology).

CRediT authorship contribution statement

Simona Sbranna: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Aviad Albert:** Writing – review & editing, Software, Methodology, Conceptualization. **Martine Grice:** Writing – review & editing, Supervision, Methodology, Conceptualization.

Acknowledgements

This research was funded by a.r.t.e.s. Graduate School for the Humanities Cologne and the German Research Foundation (DFG), grant number 281511265, SFB 1252 Prominence in Language. We thank the Goethe Institute of Naples for making rooms available for the recordings and Prof. Silvia Palermo for recruitment of participants at the University of Naples.

References

Albert, A. (2023). *A Model of Sonority Based on Pitch Intelligibility.* Studies in Laboratory Phonology 13. Berlin: Language Science Press. doi: 10.5281/zenodo.7837176.

Albert, A., Cangemi, F., Ellison, T. M., & Grice, M. (2020). ProPer: PROsodic analysis with PERiodic energy [computer software]. doi: 10.17605/OSF.IO/28EA5.

Albert, A., & Nicenboim, B. (2022). Modeling sonority in terms of pitch intelligibility with the nucleus attraction principle. *Cognitive Science, 46*(7). https://doi.org/10.1111/cogs.13161 e13161.

Albert, A., Cangemi, F., & Grice, M. (2018). Using periodic energy to enrich acoustic representations of pitch in speech: A demonstration. *Proceedings of Speech Prosody, 9*, 13–16.

Allerton, D. J. (1978). The notion of "givenness" and its relations to presupposition and to theme. *Lingua, 44*(1), 133–168.

Arvaniti, A., Ladd, D. R., & Mennen, I. (2006). Tonal association and tonal align- ment: Evidence from Greek polar questions and contrastive statements. *Language and Speech, 49*(4), 421–450.

Avesani, C. (1997). *I toni della RAI* (pp. 659–727). La Radio: Un esercizio di lettura intonativa. Gli Italiani Trasmessi.

Avesani, C., Bocci, G., & Vayra, M. (2015). Prosody and information status in Italian and German L2 intonation. In M. Chini (Ed.), *Il parlato in [italiano] L2: Aspetti pragmatici e prosodici /[Italian] L2 Spoken Discourse: Pragmatic and Prosodic Aspects* (pp. 93–116). Milano.

Avesani, C., Bocci, G., Vayra, M., & Zappoli, A. (2013). All'interfaccia tra prosodia e struttura informativa: La realizzazione prosodica dell'informazione data in tedesco e in italiano L2. In V. Galatà (Ed.), *Multimodalità e multilingualità la sfida più avanzata della comunicazione orale* (pp. 53–71). Roma: Bulzoni Editore.

[11] Studies using the continuous signal for training purposes (e.g., Gorijan et al., 2013; Smorenburg et al., 2015) have proved to be successful, despite the fact that such methods do not represent the most straightforward way to teach the results to a naive audience, nor is it the easiest to implement, as it requires training teachers first. For this purpose, a systematisation might be more practical, paying attention not to fall into simplistic generalisation. Such a systematisation requires joint work from the field of phonetics and pedagogy.

Avesani, C., Hirschberg, J., & Prieto, P. (1995). The intonational disambiguation of potentially ambiguous utterances in English, Italian and Spanish. In *Proceedings of the XIIIth International Congress of Phonetic Sciences* (Vol. 1, pp. 174–177).

Avesani, C., & Vayra, M. (2005). Accenting, deaccenting and information structure in Italian dialogue. In *6th SIGdial Workshop on Discourse and Dialogue (SIGDial 2005)* (pp. 19–24).

Barnes, J., Veilleux, N., Brugos, A., & Shattuck-Hufnagel, S. (2012). Tonal Center of Gravity: A global approach to tonal implementation in a level-based intonational phonology. *Laboratory Phonology, 3*(2), 337–383.

Baumann, S. (2005). Degrees of givenness and their prosodic marking. *International Symposium on Discourse and Prosody as a complex interface*.

Baumann, S. (2006). Information structure and prosody: Linguistic categories for spoken language annotation. *Methods in Empirical Prosody Research, 3*, 153–180.

Baumann, S., Grice, M., & Steindamm, S. (2006). Prosodic marking of focus domains – categorical or gradient. In *Proceedings of Speech Prosody* (Vol. 2006, pp. 301–304). Dresden, Germany.

Baumann, S., & Hadelich, K. (2003). Accent type and givenness: An experiment with auditory and visual priming. In *Proceedings of the 15th ICPhS, Barcelona, Spain,* 1811–1814.

Baumann, S., & Riester, A. (2013). Coreference, lexical givenness and prosody in German. *Lingua, 136*, 16–37.

Bell, A. (1984). Language style as audience design. *Language in society, 13*(2), 145–204.

Bertinetto, P. M., & Loporcaro, M. (2005). The sound pattern of standard Italian, as compared with the varieties spoken in Florence, Milan and Rome. *Journal of the International Phonetic Association, 35*(2), 131–151.

Bishop, J. (2013). Information structural expectations in the perception of prosodic prominence. In G. Elordieta & P. Prieto (Eds.), *Prosody and Meaning* (pp. 239–270). De Gruyter Mouton.

Bocci, G. (2013). *The Syntax-Prosody Interface: A cartographic perspective with evidence from Italian* (Vol. 204). John Benjamins Publishing.

Bocci, G., & Avesani, C. (2011). Phrasal prominences do not need pitch movements: Post-focal phrasal heads in Italian. In *Proceedings of the 12th Annual Conference of the International Speech Communication Association (Interspeech)* (pp. 1357–1360). Florence, Italy.

Bocci, G., & Avesani, C. (2015). Can the metrical structure of Italian motivate focus fronting? *Beyond Functional Sequence. The Cartography of Syntactic Structures (10)*, 23–41.

Boersma, P., & Weenink, D. (2021). Praat: Doing phonetics by computer [computer programme]. Version, 6.2 (Version 6.2).

Bürkner, P.-C. (2016). Brms: Bayesian regression models using Stan [R package].

Canepari, L. (1980). Italiano standard e pronunce regionali. *CLEUP.*.

Cangemi, F., Albert, A., & Grice, M. (2019). Modelling intonation: Beyond segments and tonal targets. In *Proceedings of the 19th International Congress of Phonetic Sciences* (pp. 572–576). Melbourne, Australia.

Cangemi, F., El Zarka, D., Wehrle, S., Baumann, S., & Grice, M. (2016). Speaker-specific intonational marking of narrow focus in Egyptian Arabic. In *In Proceedings of Speech Prosody* (pp. 335–339).

Cangemi, F., & Grice, M. (2016). The importance of a distributional approach to categoriality in Autosegmental-Metrical accounts of intonation. *Laboratory Phonology, 7*(1), 1–20.

Cangemi, F., Krüger, M., & Grice, M. (2015). Listener-specific perception of speaker-specific production in intonation. *Individual Differences in Speech Production and Perception*, 123–145.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software, 76*(1).

Chen, H., & Bi, R. (2008). English majors' pronunciation and intonation in read speech: A longitudinal study. *Journal of PLA University of Foreign Languages, 31*(4), 43–49.

Cole, J., & Shattuck-Hufnagel, S. (2016). New methods for prosodic transcription: Capturing variability as a source of information. *Laboratory Phonology, 7*(1).

Cruttenden, A. (1993). The de-accenting and re-accenting of repeated lexical items. *ESCA workshop on prosody*.

Cruttenden, A. (1997). *Intonation* (2nd ed.). Cambridge University Press.

Cruttenden, A. (2006). The de-accenting of given information: A cognitive universal? *Empirical Approaches To Language Typology, 20*(8), 311.

. *Tonal structure and pitch accent in Italian focus constituents, Vol. 3*, 1757–1760.

D'Imperio, M., & House, D. (1997). Perception of questions and statements in Neapolitan Italian. In G. Kokkinakis, N. Fakotakis, & E. Dermatas (Eds.), *Proceedings of Eurospeech '97* (pp. 251–254).

De Ruiter, L. E. (2015). Information status marking in spontaneous vs read speech in story-telling tasks: Evidence from intonation analysis using GToBI. *Journal of Phonetics, 48*, 29–44.

Eckman, F. R. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning, 27*(2), 315–330.

Eckman, F. R., Iverson, G. K., & Song, J. Y. (2013). The role of hypercorrection in the acquisition of L2 phonemic contrasts. *Second Language Research, 29*(3), 257–283.

Face, T. L., & D'Imperio, M. (2005). Reconsidering a focal typology: Evidence from Spanish and Italian. *Italian Journal of Linguistics/Rivista Di Linguistica, 17*(2), 271–289.

Farnetani, E., & Zmarich, C. (1997). Prominence patterns in Italian: An analysis of F0 and duration. In A. Botinis, G. Kouroupetroglou, & G. Carayannis (Eds.), *Intonation: Theory, models and applications. Proceedings of an ESCA workshop* (pp. 115–118). Athens.

Féry, C., & Kügler, F. (2008). Pitch accent scaling on given, new and focused constituents in German. *Journal of Phonetics, 36*(4), 680–703.

Féry, C., & Samek-Lodovici, V. (2006). Focus projection and prosodic prominence in nested foci. *Language, 82*(1), 131–150.

Franke, M., & Roettger, T. B. (2019). Bayesian regression modeling (for factorial designs): A tutorial. http://doi.org/10.31234/osf.io/cdxv3.

Galves, A., Garcia, J., Duarte D. & Galves, C. (2002). Sonority as a basis for rhythmic class discrimination. In *Speech Prosody 2002*, International Conference; Aix-en-Provence, France.

Gili Fivela, B., Avesani, C., Barone, M., Bocci, G., Crocco, C., D'Imperio, M., ... Sorianello, P. (2015). Intonational phonology of the regional varieties of Italian. In *Intonation in romance* (pp. 140–197). Oxford University Press.

Giordano, R. (2006). The intonation of polar questions in two central varieties of Italian. *In Proceedings of Speech Prosody*.

Gorjian, B., Hayati, A., & Pourkhoni, P. (2013). Using Praat software in teaching prosodic features to EFL learners. *Procedia-Social and Behavioral Sciences, 84*, 34–40.

Grice, M., Baumann, S., & Benzmüller, R. (2005). German intonation in autosegmental-metrical phonology. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 55–83). Oxford: Oxford University Press.

Grice, M., Ritter, S., Niemann, H., & Roettger, T. B. (2017). Integrating the discrete- ness and continuity of intonational categories. *Journal of Phonetics, 64*, 90–107.

Grice, M., Savino, M., & Refice, M. (1997). The intonation of questions in Bari Italian: Do speakers replicate their spontaneous speech when reading? *Phonus, 3*(1), e7.

Grosser, W., James, A., & Leather, J. (1997). On the acquisition of tonal and accentual features of English by Austrian learners. *Second Language Speech: Structure and Process*, 211–228.

Gut, U. (2009). *Non-native speech: A corpus-based analysis of phonological and phonetic properties of L2 English and German* (Vol. 9). Peter Lang.

Gut, U., & Pillai, S. (2014). Prosodic marking of information structure by Malaysian speakers of English. *Studies in Second Language Acquisition, 36*(2), 283–302.

Gut, U., Pillai, S., & Don, Z. M. (2013). The prosodic marking of information status in Malaysian English. *World Englishes, 32*(2), 185–197.

Halliday, M. (1967). Notes on transitivity and theme in English: Part 2. *Journal of Linguistics, 3*(2), 199–244.

Hermes, A., Becker, J., Mücke, D., Baumann, S., & Grice, M. (2008). Articulatory gestures and focus marking in German. *Proceedings of Speech Prosody*, 457–460.

Hirschberg, J., & Avesani, C. (1997). The role of prosody in disambiguating potentially ambiguous utterances in English and Italian. In *Intonation: Theory, models and applications. Proceedings of an ESCA workshop* (pp. 189–192). Athens.

Kelly, N. (2022). Change across time in L2 intonation vs segments: A longitudinal study of the English of Ole Gunnar Solskjaer. *Languages, 7*(3).

Kohler, K. J. (1991). Terminal intonation patterns in single-accent utterances of German: Phonetics, phonology and semantics. *Arbeitsberichte Des Instituts für Phonetik Und Digitale Sprachverarbeitung Der Universität Kiel (AIPUK), 25*(1), 15–185.

Krahmer, E., & Swerts, M. (2008). Perceiving focus. In C. Lee, M. Gordon, & D. Büring (Eds.). *Topic and focus. Studies in linguistics and philosophy* (Vol. 82, pp. 121–137). Dordrecht: Springer.

Ladd, D. R. (1996). *Intonational phonology.* Cambridge: Cambridge University Press.

Ladd, D. R. (2008). *Intonational phonology* (2nd ed.). Cambridge University Press.

Lepschy, A. L., & Lepschy, G. (1977). *The Italian language today.* London: Hutchinson.

Lowie, W. (2013). The CEFR and the dynamics of second language learning: Trends and challenges. *Language Learning in Higher Education, 2*(1), 17–34.

Magno-Caldognetto, E., Ferrero, F. E., Lavagnoli, C., & Vagges, K. (1978). F0 contours of statements, yes-no questions, and wh-questions of two regional varieties of Italian. *Journal of Italian Linguistics Amsterdam, 3*(1), 57–66.

Major, R. C., & Kim, E. (1996). The similarity differential rate hypothesis. *Language Learning, 46*(3), 465–496.

Mennen, I. (1998). Can language learners ever acquire the intonation of a second language? *In STiLL-speech technology in language learning, Vol. 98*, 17–20.

Möking, E., Sbranna, S., Albert, A., & Grice, M. (2024). Monitoring and assessing L2 prosodic competence. In G. De Cristofaro, B. Samu, & F. Silva (Eds.), *Language MOOCs and OERs: new trends and challenges* (pp. 159–184). Perugia Stranieri University Press.

Mücke, D., & Grice, M. (2014). The effect of focus marking on supralaryngeal articulation – is it mediated by accentuation? *Journal of Phonetics, 44*, 47–61.

Paul, H. (2013). *Mittelhochdeutsche Grammatik*. de Gruyter.

Petrov, S. V. (2021). Hypercorrection in Russian as a foreign language. In E. Bakshutova, V. Dobrova, & Y. Lopukhova (Eds.). *Humanity in the era of uncertainty* (Vol. 119, pp. 344–350). European Publisher.

Port, R., Cummins, F., & Gasser, M. (1996). A dynamic approach to rhythm in language: Toward a temporal phonology. *ArXiv Preprint*. https://doi.org/10.48550/arXiv.cmp-lg/9508007.

R Core Team. (2021). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/.

Räsänen, O., Doyle, G., & Frank, M. C. (2018). Pre-linguistic segmentation of speech into syllable-like units. *Cognition, 171*, 130–150.

Rasier, L., Caspers, J., Van Heuven, V. J., Dziubalska-Kolaczyk, K., Wrembel, M., & Kul, M. (2010). Accentual marking of information status in Dutch and French as foreign languages. Production and perception. In *Proceedings of the 6th Inter- national Symposium on the Acquisition of Second Language Speech, New Sounds 2010* (pp. 379–385).

Rasier, L., & Hiligsmann, P. (2007). Prosodic transfer from L1 to L2. Theoretical and methodological issues. *Nouveaux Cahiers de Linguistique Française, 28*(2007), 41–66.

Roessig, S., Winter, B. & Muecke, D. (2022). *Tracing the Phonetic Space of Prosodic Focus Marking.* Front. Artif. Intell., 5. Lausanne: Frontiers Media SA. ISSN 2624-8212.

Röhr, C. T., Baumann, S., & Grice, M. (2022). The influence of expectations on tonal cues to prominence. *Journal of Phonetics, 94* 101174.
Savino, M. (2012). The intonation of polar questions in Italian: Where is the rise? *Journal of the International Phonetic Association, 42*(1), 23–48.
Sbranna, S. (2023). *Developing linguistic competence in second language learning: Fluency and prosody in Italian learners of German*. Unpblished PhD Thesis available at: https://osf.io/9ca6m.
Sbranna, S., Ventura, C., Albert, A. & Grice, M. (2023). Prosodic marking of information status in Italian. *Journal of Phonetics*, 97, 101212. ISSN 0095-4470.
Selinker, L. (1972). Interlanguage. *Product Information International Review of Applied Linguistics in Language Teaching, 10*, 209–241.
Shi, S., Zhang, J., & Xie, Y. (2014). Cross-language comparison of F0 range in speakers of native Chinese, native Japanese and Chinese L2 of Japanese: Prelim- inary results of a corpus-based analysis. In *The 9th International Symposium on Chinese Spoken Language Processing* (pp. 241–244). IEEE.
Smorenburg, R., Laura. (2015). The effect of explicit training on the prosodic production of L2 sarcasm by Dutch learners of English. In *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015), Glasgow* (pp. 1–47).
Swerts, M., Avesani, C., & Krahmer, E. (1999). Reaccentuation or deaccentuation: A comparative study of Italian and Dutch. In *Proc. 14th ICPhS, San Francisco* (pp. 1541–144). Citeseer.
Swerts, M., Krahmer, E., & Avesani, C. (2002). Prosodic marking of information status in Dutch and Italian: A comparative analysis. *Journal of Phonetics, 30*(4), 629–654.
Swerts, M., & Zerbian, S. (2010). Intonational differences between L1 and L2 English in South Africa. *Phonetica, 67*(3), 127–146.
Takeda, A. (2018). *English prosodic marking of information structure by L1 Japanese second language learners*. University of Hawai'i at Manoa. PhD thesis.
Terken, J. (1984). The distribution of pitch accents in instructions as a function of discourse structure. *Language and Speech, 27*(3), 269–289.
Trofimovich, P., & Baker, W. (2006). Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition, 28*(1), 1–30.
Turco, G., Dimroth, C., & Braun, B. (2015). Prosodic and lexical marking of contrast in L2 Italian. *Second Language Research, 31*(4), 465–491.
Urbani, M. (2012). Pitch range in L1/L2 English. An analysis of F0 using LTD and linguistic measures. *Methodological Perspectives on L2 Prosody: Papers from Ml2p*, 79–83.
Verdugo, M. D. R. (2003). Non-native interlanguage intonation systems: A study based on a computerized corpus of Spanish learners of English. *ICAME Journal, 26*, 115–132.
Vietti, A. (2019). Phonological variation and change in Italian. *Oxford Research Encyclopedia of Linguistics*.
Visceglia, T., Tseng, C., Su, C., & Huang, C. (2011). Realization of English narrow focus by L1 English and L1 Taiwan Mandarin speakers. *In ICPhS*, 2074–2077.
Wagner, P. (1999). The synthesis of German contrastive focus. In *Proceedings of ICPhS* (Vol. 99).
Wenker, G. (1877). *Das rheinische Platt: den Lehrern des Rheinlandes gewidmet*. Im selbstverlage der verfassers.
Wennerstrom, A. (1998). Intonation as cohesion in academic discourse: A study of Chinese speakers of English. *Studies in Second Language Acquisition, 20*(1), 1–25.
Zimmerer, F., Jügler, J., Andreeva, B., Möbius, B., & Trouvain, J. (2014). Too cautious to vary more? A comparison of pitch variation in native and non-native productions of French and German speakers. In *In Proceedings of Speech Prosody* (pp. 1037–1041).